
Proceedings of:

Control of Precision Systems

2010 Spring Topical Meeting

April 11 to 13, 2010

*Massachusetts Institute of Technology Campus
Cambridge, Massachusetts*

The American Society for Precision Engineering (ASPE) is a multidisciplinary professional and technical society concerned with research and development, design, manufacture and measurement of high accuracy components and systems. ASPE activities encompass relevant aspects of mechanical, electronic, optical and production engineering, physics, chemistry, and computer and materials science. Membership is open to anyone interested in any aspect of precision engineering.

Founded in 1986, ASPE provides a new focus for a diverse but important community. Other professional organizations have covered aspects of precision engineering, always as a sideline to their principal goals. ASPE is based on the core of generic concepts necessary to achieve precision in any application; independent of discipline, ASPE intends to be the focus for precision technology, and to represent all facets from research to application.

*ASPE – The American Society for Precision Engineering
301 Glenwood Avenue, Suite 205, Raleigh, NC 27603
P.O. Box 10826, Raleigh, NC 27605
Telephone: (919) 839-8444 Fax (919) 839-8039
Web site: <http://www.aspe.net>*

Preface

This book comprises the proceedings of the 2010 Spring Topical Meeting. The contributions reflect the authors' opinions and are published as presented to ASPE, without change. Their inclusion in this publication does not necessarily constitute endorsement by the ASPE, or its editorial staff.

Individual readers of this book and nonprofit libraries acting for them are permitted to make fair use of the material contained in this book to be used in teaching or research; provided they copy said paper in total, as written by the author. Permission is granted to quote excerpts from the scientific or technical works in this book with proper acknowledgment of the source to include: The author's name, the book name, and the ASPE tag line. Reproduction of figures and tables is likewise permitted from the contents of this book, provided that the same acknowledgment of the same source and content is permitted.

Any other type reproduction, systematic copying or multiple reproduction of any material in this book is expressly prohibited except with the written permission of the ASPE editorial staff, addressed to the Society's headquarters.

ISBN 978-1-887706-57-706-53-7

2010

American Society for Precision Engineering
301 Glenwood Avenue, Suite 205
Raleigh, NC 27603

Printed in the United States of America

2010 Spring Topical Meeting

Contents

<i>Preface</i>	<i>ii</i>
<i>Contents</i>	<i>iii</i>
<i>Program</i>	<i>iv</i>
<i>Foreword</i>	<i>v</i>
<i>Technical Paper Index</i>	<i>vi</i>
<i>Technical Papers</i>	<i>11-150</i>
<i>Author Index</i>	<i>151</i>

2010 Spring Topical Meeting

Meeting Schedule

	Sunday, April 11		Monday, April 12		Tuesday, April 13
7:30	Registration and Continental Breakfast	7:30	Registration and Coffee	7:30	Registration and Coffee
8:30	TUTORIAL	8:15	TECHNICAL SESSION I	8:15	TECHNICAL SESSION V
10:00		10:00		10:00	
	Coffee Break		Coffee Break		Coffee Break
10:30	TUTORIAL	10:30	TECHNICAL SESSION II	10:30	TECHNICAL SESSION VI
12:00		12:15		12:00	
	Box Lunch		Lunch on Your Own		MIT Lab Tour
1:30	TUTORIAL	1:15	TECHNICAL SESSION III	1:30	TECHNICAL SESSION VII
3:00		3:00		3:00	
	Break		Break		Break
3:30	TUTORIAL	3:30	TECHNICAL SESSION IV	3:45	TECHNICAL SESSION VIII
5:00		5:00		5:00	
		5:15	Depart for MIT LIGO TOUR		
		5:30	Depart for Reception		
		5:30	RECEPTION AT LE MERIDIEN HOTEL		
		8:00			

Foreword

The third meeting on the Control of Precision Systems will focus on new developments in feedback control as applied to precision systems. We also encourage presentations of a tutorial nature that elaborate principles essential to achieving high performance in precision control systems.

The control of precision systems encompasses a wide range of applications in high-speed, high-accuracy manufacturing and automation. The growth of nanotechnology and nanoscale manufacturing has further raised the positioning requirements for machine designers. The performance of such systems, instruments, devices, and processes depends intimately upon the innovative application of feedback principles due to the large dynamic range of controlled variables and the sensitivity to effects that might be ignorable in conventional systems.

This conference is intended to promote a broader understanding of the principles and techniques applicable for precision control, to highlight the challenges and achievements unique to our field, to bring together specialists and practitioners from industry, government, and academia for the exchange of ideas, and to identify topics for further research. The conference schedule will include significant unstructured time to allow for technical and social interactions.

2010 Spring Topical Meeting Program Committee:

Meeting Co-Chairmen:

*David L. Trumper
Massachusetts Institute of Technology*

*Jeffrey W. Roblee
AMETEK – Precitech, Inc.*

*Stephen J. Ludwick
Aerotech, Inc.*

*Jan van Eijk
Delft University of Technology & MICE b.v.
The Netherlands*

Technical Sessions

Session I

Controller Architecture

Session Chairman: David L. Trumper

Monday, April 12, 2010, 8:15 am - 10:00 am

1. **A Controller Architecture to Facilitate Custom Algorithm Implementation**
Wilson, C. S. (Delta Tau Data Systems, Inc.) 11
2. **FPGA-Based Control System for Highly Dynamic Axes in Ultra-Precision Machining**
Brecher, C.; Lindemann, D.; Merz, M.; Wenzel, C. (Fraunhofer Institute for Production Technology IPT) 17
3. **Effect of Control Output Jitter on Precision Positioning**
Smeds, K.; Usman, I.; Lu, X. (The University of British Columbia) 23
4. **Tripod Floater: An "Active Magnetic Bearing" Demonstrator**
Norg, M. L. (Norg Consulting) 29

Session II

Case Study

Session Chairman: David L. Trumper

Monday, April 12, 2010, 10:30 am - 12:15 pm

1. **Case Study for an Automated Lathe Ultramicrotome**
Reck, M. L. (Boston Engineering Corp.) 35
2. **Combination of High-Speed Surface Scanning and Tactile 3-D Coordinate Measurement on the Basis of Advanced Mechatronic and Control Concepts**
Hausotte, T.; Manske, E.; Jäger, G.; Percle, B. D.; Dorozhovets, N.; Machleidt, T. (Ilmenau University of Technology) 40
3. **Precision Machine Concepts for Ultra Large Workpieces**
Bringmann, B.; Brinken, F. (StarragHeckert AG) 45
4. **High-Accuracy Atomic Force Microscope Using a Self-Sensing Probe**
Amin-Shahidi, D.; Ljubicic, D. M.; Trumper, D. L. (Massachusetts Institute of Technology); and Overcash, J. L.; Hocken, R. J. (University of North Carolina at Charlotte) 51

Session III

Feedforward, Repetitive, and Learning Control

Session Chairman: Jeffrey W. Roblee

Monday, April 12, 2010, 1:15 PM - 3:00 PM

1. **A User's Guide to Repetitive Control Systems and the Internal Model Principle**
Ludwick, S. J.; Profeta, J. A. (Aerotech, Inc.) 55
2. **Machine-Tool Tracking Error Reduction in Complex Trajectories Through Anticipatory ILC**
Madariaga, J.; Uriarte, L. G.; Ruiz de Argandoña, I.; Azpeitia, J. L. (Fundación Tekniker IK4); and Rodríguez de Yurre, J. C. (Fagor Automation S. Coop.) 59
3. **Complementary Feedback, Feedforward, and Iterative Learning Control Design for High Precision Repetitive Tracking Problems***
Bristow, D. (Missouri University of Science & Technology) 65
4. **Non-Rigid Body Feedforward for High Precision Stages***
Vervoordeldonk, M. J.; Baggen, M. (ASML) 66

* Extended abstract unavailable at press time

Session IV

Vibration Isolation

Session Chairman: Jeffrey W. Roblee

Monday, April 12, 2010, 3:30 pm – 5:00 pm

1. **The Mythical Active Vibration Isolation**
van Eijk, J. (Delft University of Technology & MICE bv); and Laro, D. A. H.; Eijsinger, J.;
Aarden, W. W. J., van den Berg, C. J. M. (MI-Partners) 67
2. **Precision Control of a Quadruple Pendulum Test Mass Suspension for the Laser Interferometer Gravitational-Wave Observatory (LIGO)**
Shapiro, B. N. (Massachusetts Institute of Technology) 69
3. **Prototyping, Testing and Performance of the Two-Stage Seismic Isolation System for Advanced LIGO Gravitational Wave Detectors**
Matichard, F.; Mittleman, R.; Barnum, S.; Foley, S.; MacInnis, M.; Mason, K.; Shoemaker, D.;
Sten, A. (Massachusetts Institute of Technology LIGO Project); Abbott, B.; Coyne, D. (California
Institute of Technology LIGO Project); Lantz, B. T.; De Bra, D. B. (Stanford University LIGO
Project); Giaime, J. A.; Hanson, J.; O'Reilly, B.; Ramet, C. (LIGO Livingston Observatory);
Gray, C.; Radkins, H. (LIGO Hanford Observatory); and Kissel, J. (Louisiana State University) 75

MIT LIGO Tour

Monday, April 12, 2010, 5:15 PM - 6:30 PM

Networking Reception

Monday, April 12, 2010, 5:30 PM - 8:00 PM

Light reception at the nearby Le Meridien Hotel. Shuttle bus will be available at the conclusion of the technical sessions and LIGO tour.

Session V

Control Design

Session Chairman: Stephen J. Ludwick

Tuesday, April 13, 2010, 8:15 am - 10:00 am

1. **Optimized Dynamic Decoupling in Scanning Stages**
Heertjes, M. F.; van Engelen, A. J. P.; Steinbuch, M. (Eindhoven University of Technology) 81
2. **Position Control and Active Damping**
Mock, C. D. (Physik Instrumente (PI) GmbH & Co. KG) 87
3. **Controller Design for Precision X-Y Stage Using Sensor/Actuator Averaging and Active Damping**
Gordon, D. J.; Erkorkmaz, K. (University of Waterloo) 92
4. **Fast Long Range Actuator (FLORA II) for Freeform Optical Surfaces**
Dow, T. A.; Garrard, K. P.; Sohn, A. (North Carolina State University) 96

Session VI

Dynamic Error Budgeting

Session Chairman: Stephen J. Ludwick

Tuesday, April 13, 2010, 10:30 AM - 12:00 PM

1. **Higher Bandwidth, Better Precision! ???**
van Eijk, J. (Delft University of Technology & MICE bv) 102

2. A Systems Approach to Modeling of Piezoresistive MEMS Sensor Panas, R. M.; Cullinan, M. A.; Culpepper, M. L. (Massachusetts Institute of Technology)	105
3. Optimal Multiple Linear Motor Commutation by Real-Time Optimization Ruben, S. D.; Tsao, T.-C. (University of California - Los Angeles)	111

Technical Tour and Lunch

MIT Laboratory Tour

Tuesday, April 13, 2010, 12:00 pm - 1:30 pm

Walking tour of several laboratories engaged in high-precision motion control research at the Massachusetts Institute of Technology. A boxed lunch will be provided to all attendees.

Session VII

Friction Modeling and Control

Session Chairman: Jan van Eijk

Tuesday, April 13, 2010, 1:30 PM - 3:00 PM

1. Fundamentals of Friction Modeling Al-Bender, F. (Katholieke Universiteit Leuven)	117
2. Practical Controller for Precision Positioning System — A New Design Approach and Application to Mechanism with Friction Chong, S.-H.; Sato, K. (Tokyo Institute of Technology)	123
3. Frequency Domain Based Feed Forward Tuning for Friction Compensation Rijlaarsdam, D. J.; van Geffen, V.; Nuij, P. W. J. M.; Steinbuch, M. (Eindhoven University of Technology); and Schoukens, J. (Vrije Universiteit Brussel)	129

Session VIII

Equipment Design

Session Chairman: Jan van Eijk

Tuesday, April 13, 2010, 3:45 PM - 5:30 PM

1. A High-Performance X-Y Stage With a Novel Topology McCarthy, K. J. (Dover); and Hunter, B. L. (Active Precision, Inc.)	135
2. Design and Control of an Over-Actuated Lightweight 450 mm Wafer Chuck Laro, D. A. H.; Boshuisen, R. (MI-Partners); and van Eijk, J. (MICE b.v. and Delft University of Technology)	141
3. A Desktop Electrohydrodynamic Jet Printing System with Integrated High-Resolution Sensing and Control Mishra, S.; Barton, K. L.; Alleyne, A. (University of Illinois)	145

Technical Papers



A S P E

A CONTROLLER ARCHITECTURE TO FACILITATE CUSTOM ALGORITHM IMPLEMENTATION

Curtis S. Wilson
Vice President of Engineering and Research
Delta Tau Data Systems, Inc.
Chatsworth, CA USA

INTRODUCTION

For many years, designers of advanced precision control systems have had to decide between the use of standard commercial controllers, which provide a full feature set but make the introduction of special algorithms difficult, or more open "prototyping" controllers (or even general-purpose computers), which easily accept special algorithms, but require the user to write algorithms himself for all aspects of the system, even where standard algorithms could suffice. These systems also tend to have a much more limited set of hardware interfaces, and lack a prioritizing scheduler.

However, the recent convergence of several hardware and software technologies enables controllers that combine the best aspects of both types of solutions. These provide a broad standard feature set, but permit the user easily and selectively to substitute or add custom algorithms using standard software tools and techniques.

ENABLING TECHNOLOGIES

Embedded Processor Advances

A new generation of embedded microprocessors has emerged recently with a feature set that makes them ideally suited for precision-control applications. As embedded CPUs, they have lower heat dissipation than most processors for general-purpose computers, permitting less extensive cooling systems.

Still, they have enough processing power to facilitate the calculations required in precision applications. In particular, many now provide hardware implementation of double-precision floating-point math, which permits fast execution of many necessary algorithms without loss of resolution or accuracy.

Software Tool Advances

Recent years have seen the widespread implementation of demanding embedded computing applications in such areas as digital

video recorders, servers, telecommunications switchers, and smart phones. Because of the ubiquity of these applications, the underlying tools, such as operating systems, have been wrung out to become very robust. Due to cost pressures, "free" open-source solutions, such as the Linux operating system, have become prominent in this class of applications.

Many of these applications have hard-real-time requirements to guarantee quick response to certain demands. As a result, hard-real-time operating systems, and kernels for general-purpose operating systems, including open-source types, have become available and robust.

At a higher level, good tools such as compilers have become readily available, inexpensive (or free!), and flexible for many important processors. Open-source versions, such as the GNU C/C++ compiler, can be incorporated seamlessly into special development environments. Modular development software tools can readily be adapted to new environments.

CONTROLLER DESIGN STRATEGY

A new motion and machine controller suitable for precision applications has been designed to take advantage of these new hardware and software tools. Delta Tau's Power PMACTM controller employs an embedded Power PC processor and the Linux operating system with a real-time kernel extension to create a fundamentally new type of controller.

The Power PMAC is a general-purpose embedded PC with a built-in motion and machine-control application. Individual features of the control application can be selected for use as desired. As a result, it can be used anywhere on the spectrum from a completely general-purpose computer to a completely dedicated controller.

Task Scheduler

At the core of the control application is a sophisticated task scheduler that sequences the many different tasks required for machine control. In many cases, the scheduler can be instructed to select between a built-in algorithm for a given task, or a custom user-written algorithm. This choice is independent for each task, so selecting a custom algorithm for one task does not dictate the choice for another task.

As a hard-real-time application, the scheduler is driven by interrupts of different priorities, many on fixed-time intervals. The priority levels, from highest to lowest, are as follows:

1. Position capture/compare list update
2. Phase commutation update
3. Servo update
4. Real-time interrupt update
5. Scheduled background cycle
6. Release to general-purpose operating system applications

Each priority level is discussed below.

Position capture/compare list update

In Power PMAC systems, the application-specific ICs that process quadrature and sinusoidal-encoder feedback have “position-capture” circuits that latch the count value (including interpolated sub-count data) immediately upon an input trigger, and “position-compare” circuits that create an output trigger immediately upon a preset count value (again including interpolated sub-count) being reached. The absence of any software delays results in very high position accuracies for the I/O.

In most cases, preparing these circuits for the next event requires software intervention, often at frequencies higher than the servo and phase updates. To support these cases, the ASICs are capable of interrupting the Power PMAC processor at the highest priority on either a capture or a compare event, triggering a user-written C routine that will take the appropriate action to prepare for the next event. Usually, this will simply be a “list update”, copying the captured position to an array, or copying the next “compare position” from an array to the active ASIC register.

Phase Commutation Update

At a frequency that is programmable, but fixed for a given application, a servo-interface ASIC in

the Power PMAC generates a clock signal that sequences the “phase” input and output circuits in the ASICs, and interrupts the processor to signal it to perform the “phase commutation” update for the motors in the system.

On each phase-clock interrupt, the processor will check each active motor in the system to see whether the user has specified phase-update tasks for that motor. If the user has, the processor then checks to see whether the user has selected the standard built-in algorithm, or a custom algorithm written by the user.

Phase algorithms compute the sign and magnitude of current required in each phase of the motor to obtain the torque requested by the servo loop, and sometimes the field strength, under certain constraints. Optionally, it can use digitized readings of actual motor-phase currents to close a current loop and output phase voltage commands, usually encoded as PWM duty cycle values.

Very-high-precision systems typically compute the phase currents for brushless servo motors in the controller and output these on high-resolution D/A-converters (16- or 18-bit) to a linearly modulated amplifier that closes an analog current loop on each phase.

Servo Update

At another frequency that is programmable, but fixed for the application (and divided down from the phase-clock signal), an ASIC generates a clock signal that sequences the “servo” input and output circuits in the ASICs, and interrupts the processor to signal it to perform the “servo” update for the motors in the system.

On each servo-clock interrupt, the processor will automatically perform a series of tasks. First, it will update the “encoder conversion table”, which processes the raw feedback data received in fixed-point format, converting it to the floating-point format the servo loop requires, with optional filtering and change-limiting.

Next, it will compute new setpoints for each active motor, updating the trajectory equations for the increment of time since the last update, and possibly computing any following of a master position sensor.

Following this, it will compute all active user-defined compensation table corrections

(position, backlash, torque) as functions of the desired positions just computed.

Finally, it will check each active motor in the system to see whether the user has selected one of the standard built-in servo algorithms, or a custom algorithm written by the user. It will execute the selected algorithm as a subroutine, using feedback and feedforward terms to compute a commanded servo control effort.

Real-Time Interrupt Update

At a programmable frequency divided down in software from the servo clock, a software “real-time interrupt” (RTI) is generated. The processor performs several tasks on this interrupt, including scans of any user foreground “PLC” programs, either in the script language, or in C.

Every RTI, the processor checks for each active “coordinate system” (collection of synchronized axes) whether a new programmed move or a move segment derived from the programmed move must be calculated. Each time a programmed move or segment finishes executing, a new one must be calculated to keep the queue full, permitting continuous motion execution.

A segment is calculated at the user-set “coarse interpolation” period (typically 10 – 20 servo cycles) for path moves, permitting time-consuming calculations to be performed at a frequency intermediate between programmed moves and servo updates. Fine interpolation at servo rates is then performed by a simple but smooth cubic-spline algorithm.

One important set of custom user algorithms that can be performed under the RTI are forward and inverse kinematic subroutines. These perform complex conversions between tool-tip coordinates and the matching underlying actuator (motor) coordinates, permitting the programming of the path motion in the user’s desired tool-tip coordinates.

The forward-kinematic routine, which converts from actuator to tool-tip coordinates, is automatically executed at the start of programmed motion to compute the initial tool-tip coordinates, needed for the first move.

The inverse-kinematic routine, which converts from tool-tip to actuator coordinates, is automatically executed for each programmed

move (or each derived segment for a path move) to compute the actuator positions necessary to implement the programmed tool-tip motion.

Background Cycle

In time left over between interrupt tasks, the processor cycles through user background programs, which can be both in the provided script language and in C. As background tasks, the execution of these programs is not deterministic, limiting their utility for many control applications, but these can still be used for many less time-critical algorithms.

Operating System Applications

The controller’s scheduler periodically releases from background tasks to permit other applications to execute directly from the Linux operating system. These applications execute in the general-purpose operating system, so their execution is not deterministic.

INTEGRATED DEVELOPMENT ENVIRONMENT

A key tool for implementing an application with the Power PMAC is the Integrated Development Environment (IDE) software on the PC. This program facilitates the creation and management of the entire Power PMAC application.

The IDE includes many components comparable to those for general-purpose software development (plus others specific to motion and machine control). Key components for custom algorithm generation include a project manager, an advanced editor, and an embedded cross-compiler.

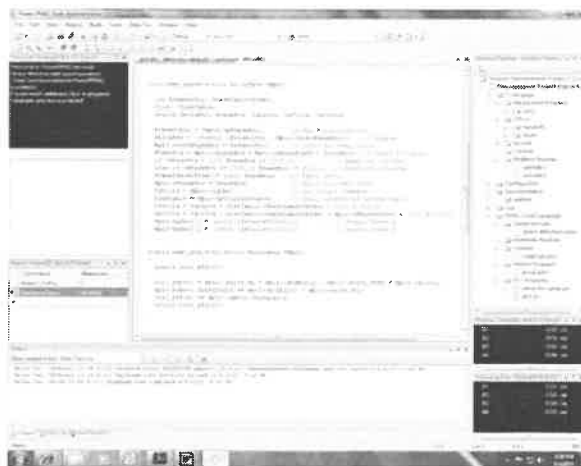


FIGURE 1. The Power PMAC IDE Program

Project Manager

The project manager in the IDE permits the user to manage the different software components of the application project in an organized manner. The programs, routines and files are organized into different folders according to function. Project files are stored in these folders on both the PC and the controller.

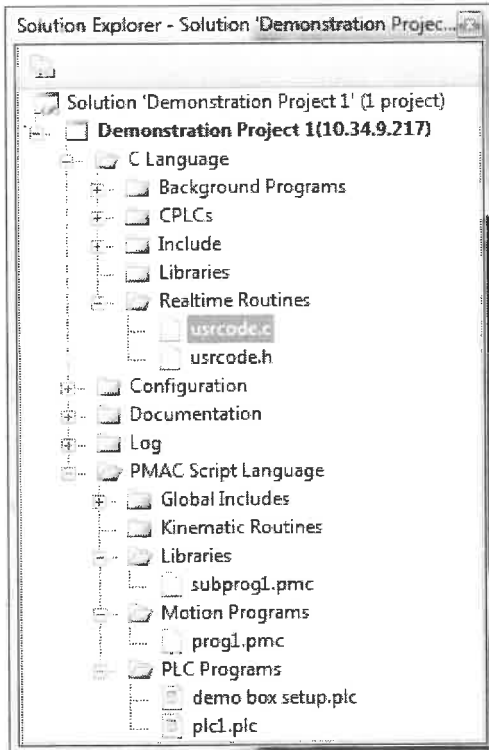


FIGURE 2. The IDE Project Manager

When you create a new file in a given folder, it is automatically started with a template appropriate to the type of program or routine for that folder.

Editor

The IDE's editor has many advanced functions familiar to modern software development. For both the Power PMAC script language and C/C++, it provides automatic color coding, syntax checking, and parentheses matching.

It also provides an "intellisense" feature, automatically providing relevant completions for language reserved words (both script and C), pre-defined data structure element names, and user-declared variable names.

For script-language reserved words, and pre-defined data structure element names, the F1

function key automatically presents a detailed description from the Software Reference manual.

Cross-Compiler

The IDE embeds within it the open-source GNU C/C++ cross-compiler for the Power PMAC's processor. A simple command to "build" the project, or part of a project, will invoke the compiler to process the C/C++ source code in the relevant files, and load the resulting machine-code files into the Power PMAC. The process is as simple for the user as loading interpreted script code.

Plotting and Tuning Utility

The Power PMAC can gather the contents of any set of registers at the servo or phase update rates and pass the resulting data file to the IDE. This can be for automatic tuning sequences such as steps and ramps, or for user-defined sequences. The IDE's plotting utility provides a powerful and flexible graphic display of the resulting data.

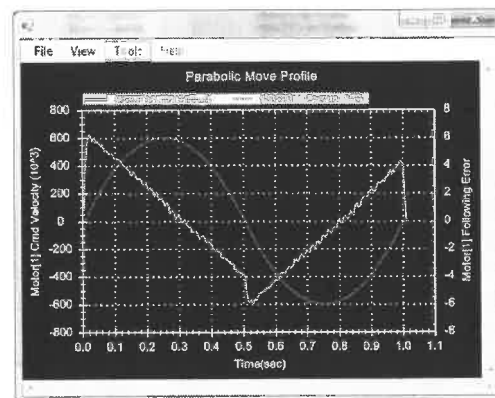


FIGURE 3. Sample IDE Plotted Data Display

WRITING CUSTOM PHASE AND SERVO ALGORITHMS

Custom algorithms to perform the commutation and servo-loop tasks can be written in C or C++. When written in the proper file and folder, a project "build" will automatically compile and download the resulting executable code to the proper location in the Power PMAC.

A given motor may use a custom phase routine, a custom servo routine, both, or neither. For both servo and phase, multiple custom routines may be used in the same application. Multiple motors may use the same custom algorithm. Other motors may use the standard algorithm.

Note that the algorithm “slot” does not have to be used for actual phase or servo tasks.

The algorithm file must include the directive:

```
#include <RtGpShm.h>
```

to be able to access the pre-defined Power PMAC data structure elements in shared memory. Each motor structure has 32 64-bit registers reserved for custom user algorithms.

Custom Phase Commutation Algorithms

A custom algorithm to perform commutation of multi-phase motors, and/or close a digital current loop, must be declared in the form:

```
void MyPhaseAlg(struct MotorData *Mptr)
```

The argument provides a pointer to the present motor, giving the algorithm direct access to that motor’s complete data structure. The algorithm does not need to know which motor is calling the routine, so it can be made motor-independent and callable by multiple motors.

The algorithm can use data structure elements that would normally be used by the standard algorithm. This is strongly encouraged, especially for saved setup elements such as the input and output pointers.

The algorithm is likely to start with a statement such as:

```
PresentPos = (double) *Mptr->pPhaseEnc;
```

which will read the value at the address of the hardware register specified by the pointer element **pPhaseEnc**, and convert the fixed-point value to double-precision floating-point format recommended for the actual calculations. The standard algorithm uses this element to specify the address of the rotor angle feedback value.

The commutation subroutine does not return a value (hence the **void** declaration), so it is necessary for the algorithm to write directly to the output registers. The routine is likely to finish with statements such as:

```
Mptr->pDac[0] = (int) PhaseUCmd;  
Mptr->pDac[1] = (int) PhaseVCmd;
```

These statements convert the floating-point phase command values to integer format and

write them to two consecutive addresses specified by the saved pointer element **pDac**.

The algorithm has access to Power PMAC’s built-in 2048-entry sine table, with each entry a double-precision floating-point value. This eliminates the need for time-consuming trigonometric computations.

Custom Servo Algorithms

A custom servo algorithm must be declared in the form:

```
double MyServoAlg(struct MotorData *Mptr)
```

As with the phase, the argument provides a pointer to the present motor for access to that motor’s complete data structure.

The servo-update processes manage the inputs and outputs for the servo algorithm more than for the phase. The user’s algorithm does not need to use pointers for these; the already processed floating-point inputs are available as:

DesPos: Net commanded position

ActPos: Net actual position

ActPos2: (Optional) 2nd actual position

PosError: Servo error (Des – Act)

ActVel: Change in 2nd actual position

The servo command is the returned value of the routine. It will either be used as the torque (quadrature current) input to the motor’s commutation algorithm, or converted to fixed-point format and output.

Automatic routines manage the state of the servo loop: enabled vs. disabled, and closed vs. open-loop. However, it is the user’s responsibility to take the appropriate action in the algorithm based on the present state. A typical routine will take the general form:

```
if (Mptr->ClosedLoop) {  
    [Compute ControlEffort]  
}  
else {  
    [Clear integrators,ControlEffort]  
}  
return ControlEffort;
```

MIMO Algorithms

Multiple-input/multiple-output (MIMO) algorithms can be written for both servo and phase routines. For these, some code must be made specific to the motor number.

Using Matlab/Simulink™

The Mathworks' Matlab/Simulink™ software provides a powerful and flexible graphical programming environment for algorithms such as these. The Power PMAC and its IDE are designed to make implementations of algorithms designed in this environment very easy.

The user starts by designing the algorithm generically in Simulink, often with a simulated plant model and excitation signal. When ready to try the algorithm in Power PMAC, the plant transfer function and any excitation blocks should be removed and replaced with provided Power PMAC input and output blocks. These can be obtained from the "DELTATAU PPMAC Library", which is available at no charge and must be put in the "MATLAB" path of the Simulink Library Browser. The I/O blocks permit the user to specify read and write access to data structure elements such as command and actual positions.

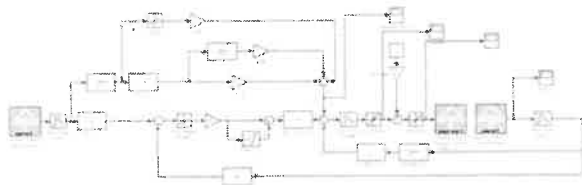


FIGURE 4. Simulink servo algorithm with Power PMAC input/output blocks (shaded)

C source code can then be generated automatically by invoking the Real-Time Workshop™ "Build Model" action under the Simulation menu. Prior to this, a one-time setup of the "Configuration Parameters" will need to be done under the same menu, so the generated code will be compatible with the Power PMAC. The results of this generation will be stored to files.

The resulting ".c" source code and ".h" header files are then brought into the "C Language/Realtime Routines" folder of the IDE's Project Manager. Manual modification is possible at this point. From here, a project "build" will compile this code and load it into the Power PMAC.

WRITING CUSTOM KINEMATIC ROUTINES

Forward and inverse-kinematic routines for the Power PMAC are written in the controller's script language, which is easier to use for non-

specialists than C is (due to features like automatic type-matching of variables), but provides sufficient power and flexibility for these algorithms.

The forward-kinematic routine for a coordinate system can expect to find the present commanded actuator (motor) positions in specified registers that can be accessed by the variable names **KinPosMotor n** , where " n " is the motor number. It must leave the calculated tool-tip (axis) coordinates in similar specified registers that can be accessed by the variable names **KinPosAxis α** , where " α " is the letter designator for the axis. The standard motion calculations will automatically use these values as the starting positions.

The conceptual form for a 2-axis forward-kinematic routine using Motors 1 & 2 and the X & Y-axes is:

```
open forward
KinPosAxisX =  $f_1$ (KinPosMotor1, KinPosMotor2);
KinPosAxisY =  $f_2$ (KinPosMotor1, KinPosMotor2);
close
```

The inverse-kinematic routine works in reverse fashion. It can expect to find commanded axis positions in the **KinPosAxis α** variables, and it must leave the calculated motor positions in the **KinPosMotor n** variables.

The conceptual form for the inverse-kinematic routine for the same 2-axis system is:

```
open inverse
KinPosMotor1 =  $g_1$ (KinPosAxisX, KinPosAxisY);
KinPosMotor2 =  $g_2$ (KinPosAxisX, KinPosAxisY);
close
```

Of course, the "functions" in both routines can be interrelated complex combinations of math and logic. Subroutines can be called. Iterative solutions are permitted for cases where no closed-form solution exists.

CONCLUSIONS

The modular software architecture of the Power PMAC provides significant benefits to users who wish to implement non-standard algorithms in some aspect of a precision motion-control system. Creation and integration of such algorithms are supported with modern software development tools. Robust standard algorithms are available for all other aspects of the control.

FPGA-BASED CONTROL SYSTEM FOR HIGHLY DYNAMIC AXES IN ULTRA-PRECISION MACHINING

Christian Brecher¹, Dominik Lindemann¹, Michael Merz¹, Christian Wenzel¹

¹Department of Production Machines

Fraunhofer Institute for Production Technology IPT
Aachen, Germany

ABSTRACT

The developments presented in this paper focus on the implementation of an FPGA-based (Field Programmable Gate Array) hybrid servo control system for mechanically coupled highly-dynamic axes. Consisting of an FPGA-based low level part for high bandwidth control loops and a microprocessor-based high level part for setpoint generation, the novel servo system enables the control of highly dynamic axis needed in Fast Tool assisted ultra-precision machining. Movements with a bandwidth of up to 100 Hz and a stroke of several mm are generated by an electro-magnetically driven axis. For highly dynamic movements with a bandwidth of up to 2 kHz a piezo-actuator is used. The presented system realizes position control clocks of above 200 kHz and thereby strongly improves the performance of high-bandwidth axes.

INTRODUCTION

Complex optical components more and more become the essential functional elements in advanced products. Examples range from miniaturized illumination systems for display technology in mobile phones or notebooks over technical optics for LED headlights in the automotive industry to bifocal glasses or contact lenses in medical applications. Replicating advanced optics by injection molding offers an economical method for the production of such freeform components. With rising geometrical complexity and required surface quality, the manufacturing of high-end optical parts and molds by diamond machining demands highly dynamic and ultra-precise axes [1]. Ultra-precision lathes, additionally equipped with a Fast Tool Servo (FTS) system, allow for the machining of non-rotationally symmetric elements with a very low surface roughness. In the turning process very accurate positioning as well as a highly dynamic movement of the cutting tool are essential to achieve the aimed surface quality [2]. Due to the need of higher control and PWM clocks such highly dynamic mechatronic axis systems cannot be con-

trolled with standard servo control systems. Thus, innovative, high bandwidth feedback control loops have to be developed. Additionally, the control system should be capable of compensating systematic positioning errors. The synchronization with the control system of the basis machine further needs an interpolation error minimized setpoint generation to achieve the required precision. To meet these demands the presented hybrid servo control system is developed based on an FPGA controlling an amplifier provided with a PWM clock of up to 900 kHz.

FAST TOOL ASSISTED TURNING

Manufacturing non-rotationally symmetric components with a very low surface roughness can be realized on ultra-precision lathes with an additional Fast Tool system. The mechanical design of a hybrid FTS system is shown in *FIGURE 1*. The piezo-driven axis (Fast Tool) with a stroke of 30 μm holding the diamond tool is supported by an electro-magnetically driven axis (Slow Tool) with a stroke of 25 mm. Both axes move independently and have their own guiding and measurement systems.

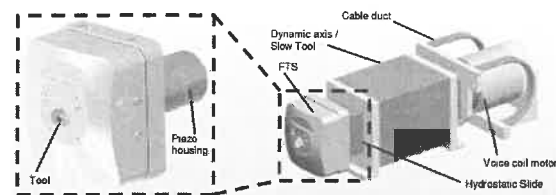


FIGURE 1. Hybrid Fast Tool Servo

While the piezo axis has a limited stroke, non-rotationally symmetric freeform parts in the range of millimeters can be generated using the electro-magnetic axis, which is driven by a voice-coil actuator. Voice-coils generate no electro-magnetic forces perpendicular to the direction the motor is intended to work. Thus, the drive exerts no significant forces onto the bearings which would influence the vertical position of the axis and decrease the machining accu-

racy. The axis is supported by hydrostatic bearings featuring only dynamic and no static friction. Capacitive sensors are applied for the position feedback. The stroke of such a system is basically limited by the combination of the piezo actuator and the stiffness of the flexural guide used for the cutting tool. Both act as springs, that work in a serial arrangement, and in that way limit the remaining stroke. To achieve maximum dynamics the properties of the guiding and the piezo actuator have to be adjusted thoroughly.

CONTROL LOOPS AND COMPENSATION

FIGURE 2 shows the information flow of the fundamental control signals between the FPGA-based servo control system and the electromechanical axes containing sensors and actors. The position control loop for the piezo axis is closed over a voltage controlled output stage for the actuator and a capacitive position sensor. It uses a control clock of 250 kHz. The transfer behavior of the electro-mechanical system is limited to 2 kHz due to mechanical eigenfrequencies. A sampling control demands at least the factor 10 as minimum limit for the control clock. To utilize the full bandwidth of the mechanics, a minimum control clock of 20 kHz is required. Though, the desired control clock is specified higher to provide a phase decay, which is as low as possible in the range of higher frequencies. This phase decay is critical for the performance of the transfer behavior of the axis.

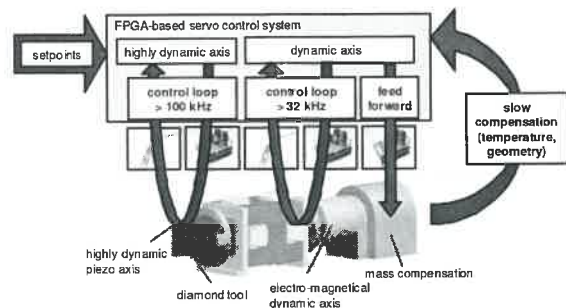


FIGURE 2. Information flow and control loops

Due to the higher mass, which has to be moved, the electro-mechanical system has clearly lower dynamics than the piezo-based system. Especially for the velocity control loop for the implemented system a clock of higher than 32 kHz is desired. Disturbances caused by the mechanical coupling of both axes can be reduced by a control oriented feed-forward decoupling, which is implemented in parallel to the feedback controls of the axes. The control clocks should be cho-

sen multiples of each other to facilitate such an implementation. Compensation strategies, which are slow in terms of the mentioned control clocks, for geometrical or thermal deviations can be implemented with a significantly slower clock than the control loops. Depending on the mathematical complexity, these algorithms are implemented in the microprocessor-based part of the system.

Servo Control Algorithms

To close the position control loop in the servo amplifier, the cascaded P/PI control algorithm, depicted in FIGURE 3, is applied. The input signal of the position controller is the position error, which is the difference between the interpolated position measurement and the position setpoint. The feed-forward control branch of the setpoint adjusts the amplitude and the phase of the control system by means of the inverse transfer function of the closed loop [3].

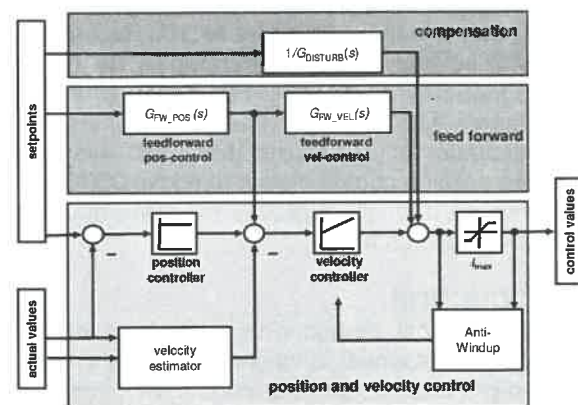


FIGURE 3. Servo control algorithm

In the case of the Slow Tool control, the velocity estimator is realized by an observer structure, which computes the actual velocity from the acceleration proportional electrical current of the voice coil and the actual position from the interpolated glass scale signal [4]. A velocity signal can either be calculated by integrating an acceleration or by differentiating a position signal. While the dominant error resulting from the integration is in the low and quasi-static frequencies, differentiations turn out to be noisy in the high frequencies. The observer mixes the signal in the frequency domain. A time constant determines a cut-off frequency, which separates the frequency ranges for the acceleration and the position signal with a transition frequency range for a smooth switching. FIGURE 4 shows the improvement achieved by the velocity observer.

The velocity gained from just differentiating the position signal contains much more noise than the velocity signal obtained by the observer. The drawback of the observer is the additional phase decay visual in the figure as a delay in the signal. Even so, the dynamic effect onto the position control bandwidth is tolerable for the Slow Tool system. The velocity error is calculated by subtracting the actual velocity estimated by the observer from the output signal of the position controller and serves as the input signal of the PI velocity controller. The output signal of the velocity controller is restricted to the output range of the power amplifier's input signal. The Anti-Windup circuit detects a saturation of this signal and inhibits further increases in the integral part of the velocity controller, which would lead to instabilities of the control loop. To prevent an algebraic loop, the calculation of the Anti-Windup is delayed for one control clock.

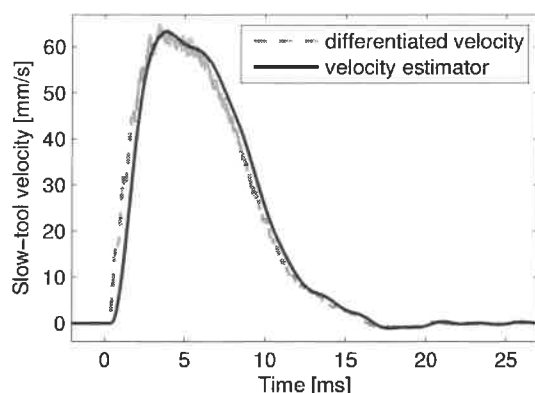


FIGURE 4. Differentiated and estimated velocity

Control Parameters and Filter Design

In addition to the standard PI controller a notch controller was tested. FIGURE 5 shows the position transfer behavior for small signals for the different configurations. The cut-off frequency as criteria, the notch control has no clear advantage. At least the amplitude curve shows a better performance than the standard PI control. Applying the feed-forward control filter, the cut-off frequency, which is defined as the frequency at the phase crossover at -90° , is increased from around 450 Hz to around 1 kHz. Up to 2 kHz the measurement shows a good coherence of the input and output signals. The bad coherence beyond 2 kHz results from the mechanic eigenfrequencies in this range. FIGURE 6 shows the derivation of the feed-forward control filter with the intermediate data needed. The shown frequency response is measured

using a sine signal with a linearly ascending frequency and a hyperbolically descending amplitude. This signal stimulates all frequencies in the desired bandwidth.

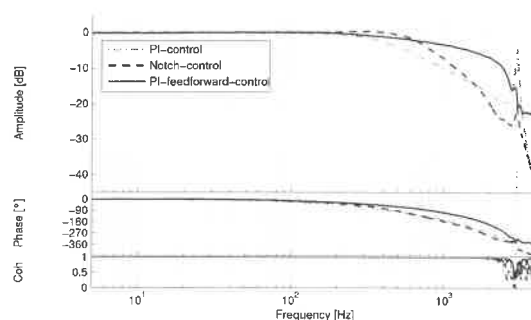


FIGURE 5. Fast Tool: position transfer behavior

The measurement is transformed from the time domain to the frequency domain with the aid of the H1/H2-method [5] for transfer behavior, which bases on the power density spectrum of the input and the output signal. While the H1-method minimizes noise on the input, with the applied H2-method the output noise is minimal. The coherence signal shows the significance of the measurement regarding the correlation of the input and output signals.

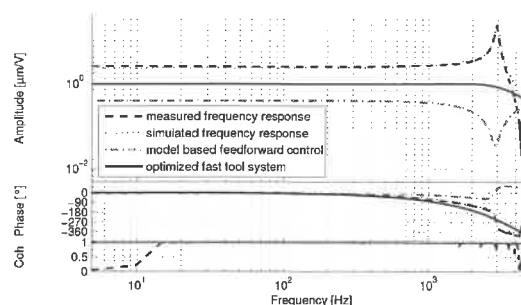


FIGURE 6. Fast Tool: feed-forward control filter

The model-based feed-forward control filter is derived from the simulation model of the transfer behavior of the closed loop system by inverting the model and adding a low pass filter with a sufficient order. The resulting filter is shown in FIGURE 6. The optimized transfer behavior of the response to position setpoint changes with the filter applied clearly shows better amplitude characteristics. Due to the low pass filter needed for the feed-forward control, the phase has a higher decay in the range of the resonance frequency compared to the original system. The disturbance reaction of the control loop is not affected by this extension.

PLATFORM ARCHITECTURE

The basic platform for the servo control system presented in the paper is a hybrid system, consisting of a microprocessor-based high level part and an FPGA-based low level part. Logically higher application layers are implemented in the microprocessor-based part of the control system. These layers include for instance parameterization and supervision interfaces to the Human Machine Interface (HMI), which do not need to be real-time capable. Further, low bandwidth control loops for compensation purposes, which might be computationally demanding are implemented in the high level part. The high bandwidth position control loops of the axes are implemented in the low latency FPGA-based part of the servo system to achieve the best possible dynamics and accuracy. The control system, which implements the control algorithms in this setup, does not constrain or limit the dynamic performance of the overall feedback system consisting furthermore of power electronics, mechanics and sensors.

In a microprocessor system the digital hardware structure of the computation system is fixed. Hence, in such a system the subset of available elementary arithmetic operations usually is restricted. In addition, the system peripherals, to which the sensors and actors are connected, in most microprocessor architectures are not connected to the arithmetical unit directly, but over a peripheral bus system. This implies a significantly high latency time from the input to the output signals, which dominates the dynamic performance of the control loops. Beyond that, the latency is not constant in time, which, if not taken account of, causes additional noise in the control loop, leading to a poor positioning performance.

Compared to a microprocessor, the digital hardware structure of a computation system, which can be implemented within an FPGA, is freely programmable and reconfigurable. The hardware structure can be adapted to the application. Especially, critical paths in the digital algorithms can be optimized. While programs on a microprocessor-based system are as a matter of principle sequentially processed, on an FPGA-based system calculations, which are adjusted to the application, can be executed in parallel. Parallelizations of this kind are especially effective in the case of modules, which are directly communicating with the electronic peripherals of the system, like measurement interfaces or con-

trol value outputs. With these techniques the overall latency of the control algorithms can be optimized to a minimum. However, while it is technically possible, FPGAs are not suited for the implementation of mathematically complex algorithms, in contrast to microprocessors with highly clocked digital logic, which is exceedingly optimized for these tasks. The implied performance discrepancy particularly applies for algorithms using floating point arithmetic. For this reason, the mathematically complex algorithms for the calculation of surface-based setpoints are processed in the microprocessor-based part of the system. On the other hand, the mathematically elementary feedback control algorithms for the highly dynamic axes benefit from the achievable extremely low latency of the FPGA.

Mathematics on FPGA

Compared to fixed point arithmetic, which can be easily implemented in an FPGA, floating point arithmetic generally achieves a better signal-to-noise ratio. While the precision of fixed point numbers is constant over the whole number range, the precision of the floating point representation depends on the size of the actual number and is distributed roughly logarithmically over the number range. This plainly means that small numbers have a very high precision, while large numbers are distributed broader. The accuracy of the mathematical calculations in a specific range is automatically adapted to the actual size of the numbers. This property facilitates the implementation of control algorithms.

Today's microprocessors generally provide hardware support for floating point operations. However, the support for floating point operations on the basis of an FPGA is at least rudimentary. Fully parallel designs can be implemented easily by just instantiating the needed number of floating point operation units provided as macro blocks by the FPGA manufacturer. Using this approach, the available capacity of the FPGA is rapidly reached even for small algorithms. On the other hand soft processors like the Xilinx MicroBlaze support floating point operations, but on these processor structures floating point operations can only be sequentially processed. Graphical programming platforms like Matlab/Simulink with the appropriate FPGA programming extension can be used to implement fixed point algorithms very fast, but do not support floating point operations at all.

To close the gap between the strictly parallel and the strictly sequential approaches, a VLIW (Very Long Instruction Word) -based processor architecture was implemented, which allows for the parallel execution of a floating point addition, a floating point multiplication, a register assignment and a flow control operation. A simplified block diagram is shown in *FIGURE 7*.

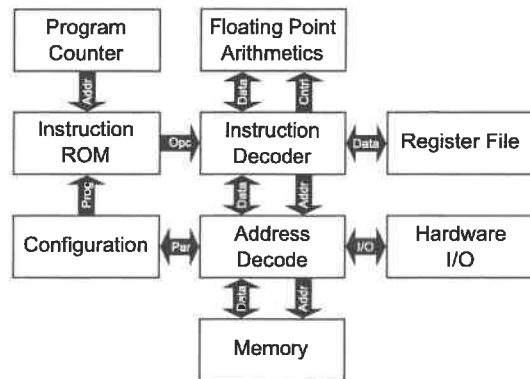


FIGURE 7. VLIW floating point processor

The processor design is easily scalable, regarding the number and type of floating point operation units as well as the number of registers. Central components of the processor are the *instruction decoder* with the *register file*, consisting of 7 registers with a width of 32 Bit each. The *instruction decoder* in each clock cycle receives a VLIW opcode (Opc) from the *instruction ROM*, which is addressed by the *program counter*. The *instruction decoder* controls the *floating point arithmetics* and the *address decoder*, which manages the access of the local memory, the hardware I/O, and the configuration containing control parameters and flags.

The program can be loaded into the *instruction ROM* over the configuration interface. The *instruction decoder* is light weight and just fills the pipelines of the *floating point arithmetics* unit and the memory operation unit with commands. Notably, it has no capability to stall operations. Thus, especially the sequence of operations needs to be generated by the compiler.

Implementation of Servo Control System

FIGURE 8 shows the actual implementation of the servo control system. The microprocessor-based high level layer is realized by an embedded control PC, which is connected to the HMI PC of the basis machine control. The low level FPGA-based part of the system is connected to the highly dynamic axes and the real-time con-

trol system of the basis machine control. The FPGA system implementation can be debugged over JTAG (Joint Test Action Group – IEEE 1149.1), the common method to debug ICs (Integrated Circuit) and embedded systems.

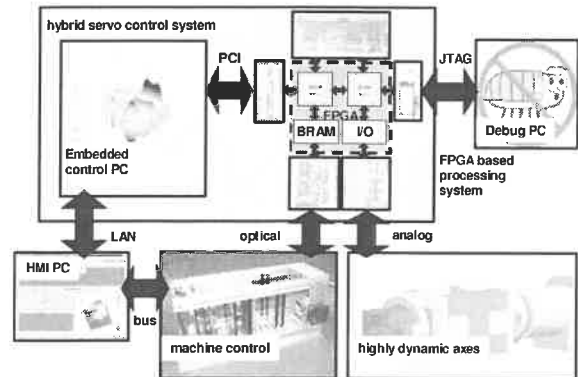


FIGURE 8. Servo control and its interfaces

Machining Experiments / Measured Results

The standard deviation of the position resolution in quiescent state is 2 nm for the Slow Tool and 7 nm for the Fast Tool (see *FIGURE 9*, right). This value is a good measure for the noise in the system accounting for the resulting surface roughness of the work piece. Machining experiments with the introduced control parameters show that the surface roughness on the work piece with 5 nm Ra is in the range of the values measured in the control system. The figures on the left show the step responses of the two dynamic axes. It is clearly visible that the feed-forward control filter improves the positioning behavior. Regarding steps, the Fast Tool system, due to the lower damping capability of the mechanical design, is more susceptible to vibrations caused by mechanical eigenfrequencies.

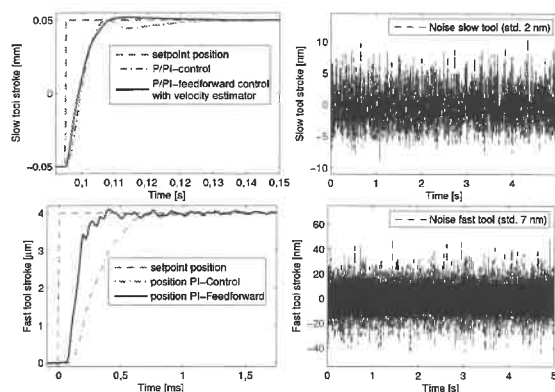


FIGURE 9. Step response and position noise

Comparing Microprocessor and FPGA

The control loops benefit significantly from the achievable control clocks using the FPGA-based control system. *FIGURE 10* shows measurements of the step response on 2 μm and the small signal transmission behavior of the Fast Tool system using the two different control platforms with a standard PI controller for comparability. The prototyping system used as reference with an architecture based on a microprocessor (Processor in *FIGURE 10*) features a sample rate of 40 kHz. The FPGA control loop is clocked with 250 kHz.

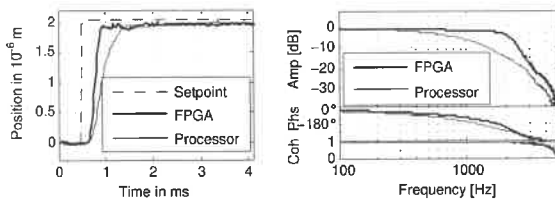


FIGURE 10. FPGA vs. microprocessor

In the time domain the response time decreases from 400 μs to 300 μs , where a significant amount of around 200 μs arises from the time response of the applied piezo amplifier and the signal processing of the capacitive sensor. As shown in *FIGURE 10*, the bandwidth specified by the edge frequency of the phase at -90° is increased from around 600 Hz to 900 Hz, with a gain in phase achieved of 40° at 900 Hz. The bandwidth regarding the -3 dB criterion is even more improved from around 600 Hz to 2 kHz.

For reasons of structure complexity yet applicable in the used Xilinx Virtex 4 FPGA only Single Precision Floating Point (IEEE754) is used, while Double Precision Floating Point is common in modern microprocessors. For feedback control loops, nevertheless, the advantages of the FPGA-based control platform regarding the achievable extremely low latency heavily outweigh the disadvantages regarding computing power, as can be seen in the measurement results presented in the paper.

CONCLUSIONS

Feedback control clocks of up to 20 kHz as applied to conventional servo drives are not capable to be used for the control of highly dynamic axis such as Fast Tools. The FPGA-based approach presented in this paper guarantees the full use of the mechanical bandwidth. Accompanied with a microprocessor-based prototyping system, where control algorithms can be tested

in principle before being implemented within the FPGA, the system is even capable for rapid control prototyping. Realizing control loops clocked with 250 kHz, the dynamics of high-bandwidth axes can be significantly improved. Comparisons to standard servo controls show promising performance improvements achieved by the FPGA-based hybrid control system.

ACKNOWLEDGEMENTS

The research work focusing on increasing the positioning accuracy of highly dynamic axes presented in this paper has been conducted in the SFB/TR4 "Process Chains for the Replication of Complex Optical Elements", funded by the German Science Foundation DFG.

The implementation of the electro-mechanical axis system and the control hardware has been conducted in the project »ERANET-OPTICALSTRUCT« (02PG2623) funded by the German Federal Ministry of Education and Research (BMBF) and managed by the project management agency of Karlsruhe Institute of Technology (PTKA-PFT).

REFERENCES

- [1] C. Brecher, M. Merz: Freiformoptiken für die Beleuchtungstechnik von Morgen – Ultrapräzisionstechnik mit Fast-Tool-Servo-Systemen, Nordrhein-Westfälische Akademie der Wissenschaften, Ferdinand Schöningh Verlag, Vorträge I 23, ISBN 978-3-506-76589-5, 2008
- [2] C. Brecher, M. Merz, C. Wenzel: Optimisation of Servo Control for Highly Dynamic Axes for Ultra Precision Freeform Machining, PCIM Europe, Nürnberg, 2007
- [3] C. Brecher, S. Lange, M. Merz, F. Niehaus, C. Wenzel, M. Winterschladen, M. Weck: NURBS Based Ultra-Precision Free-Form Machining, CIRP, Kobe (JP), 2006, pp 547-550.
- [4] C. Brecher, M. Merz, F. Niehaus, C. Richterich, K. Rinko, E. Schalle, K. Schmidt, M. Stautner, A. Ventura-Traveset, M. Zavelberg: Machine and Process Development for the Robust Machining of Micro-structures on Free-Form Surfaces for Hybrid Optics", Final Report of BMBF collaboration project ERANET-OpticalStruct, 2009, ISBN 978-3-940565-79-2.
- [5] M. Weck, Werkzeugmaschinen 5: Messtechnische Untersuchung und Beurteilung, 6. Auflage, Springer Verlag, 2006, ISBN 3-540-67615-5, pp 228-261.

EFFECT OF CONTROL OUTPUT JITTER ON PRECISION POSTIONING

Kristofer Smeds¹, Irfan-ur-rab Usman¹, Xiaodong Lu¹

¹The University of British Columbia
Vancouver, British Columbia, Canada

INTRODUCTION

The block diagram of a typical digital control system including control output disturbance, $d(t)$ and output noise, $n(t)$ is shown in Figure 1. The pure delay, τ_d , represents the controller computation time. The control input sampling element and the control output zero-order-hold element are usually considered to be perfectly synchronized with a constant sampling period of T_0 ; however, in reality there exists some deviation between the ideal timing and the actual timing. On the control output this deviation is referred to as control output jitter, $\tau_j(k)$. It is caused by factors such as task scheduling, interrupt handling, and resource sharing.

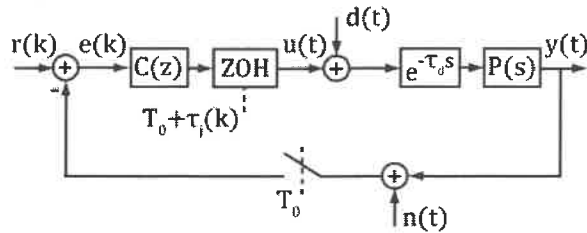


FIGURE 1. A typical digital control system block diagram.

The presence of control output jitter introduces additional positioning error into a system. Generally, this effect is considered to be negligible because in most cases jitter is small. Control output jitter for a general purpose high-end commercial control platform, such as a dSPACE 1103, is measured to be only 140ns RMS. A very well designed digital control platform [1] may exhibit control output jitter as low as 5ns RMS. It is only when dealing with precision high speed motion control systems that the effect of control output jitter can become significant. High speed systems are more susceptible because the magnitude of the jitter effect is inversely proportional to the sampling period. The additional positioning error may still be negligible unless other disturbance sources in the system have been reduced to the point that the jitter effect becomes the dominant source of error, which can be the case for precision high speed systems.

In [2], control output jitter is considered to be a positive-only variable delay and for analysis is treated as a constant gain. This approach fails to capture jitter's interaction with the measurement noise and reference commands that results in added positioning error. Here, these interactions are investigated and a detailed analysis is carried out to predict the resulting regulation and tracking error. Further, several solutions are presented to mitigate the effect of jitter. To validate the analysis a simulation of a precision magnetic bearing spindle is performed.

MODELLING OF CONTROL OUTPUT JITTER

Control output jitter is defined as the difference between the actual control output timing and the ideal control output timing. It is a zero bias signal because any DC component is included as part of the pure delay element, τ_d . For an arbitrary input, the ideal control output signal with zero jitter, $u^*(t)$, is compared to the actual signal, $u(t)$, in Figure 2.

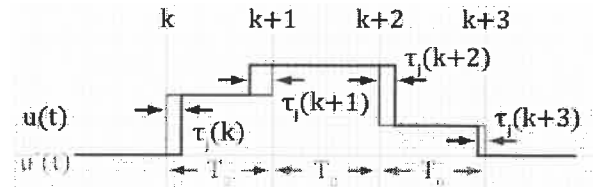


FIGURE 2. Effect of jitter on the control output. $u(t)$ is the actual control output signal and $u^*(t)$ is the ideal control output with zero jitter.

The resulting disturbance on the control output signal due to jitter is

$$g(t) = u(t) - u^*(t) \quad (1)$$

Alternately, this can be expressed as

$$g(t) = u_{\Delta}(t)j(t) \quad (2)$$

where

$$u_{\Delta}(t) = u^*(k-1) - u^*(k) = u^*(k)(z^{-1} - 1) \quad (3)$$

$$\text{when } \tau_j(k) \geq 0; j(t) = \begin{cases} 1 & t \in [kT_0, kT_0 + \tau_j(k)] \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$\text{when } \tau_j(k) < 0; j(t) = \begin{cases} -1 & t \in [kT_0 + \tau_j(k), kT_0] \\ 0 & \text{otherwise} \end{cases}$$

The above analysis results in the block diagram shown in Figure 3. The effect of jitter has now been included as a disturbance, $g(t)$, on the control output. The form of this block diagram is similar to that from [2]; however, the selection for $j(t)$ used here results in jitter being modeled as an input that modulates with the derivative of the discrete control output.

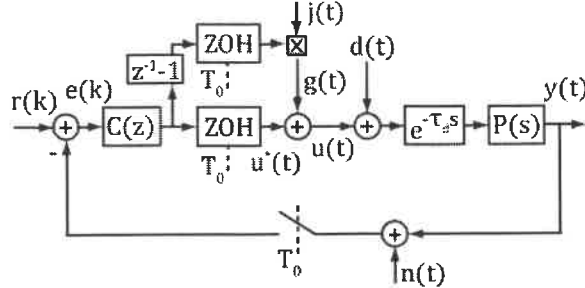


FIGURE 3. Hybrid model of control output jitter.

Based on Figure 3, this paper presents a novel method which models the effect of control output jitter as a discrete modulation by approximating $j(t)$ as the discrete dimensionless control output jitter, $\lambda(k)$. This can be expressed as

$$\lambda(k) = \frac{\tau_j(k)}{T_0} \quad (5)$$

This selection for $\lambda(k)$ conserves the disturbance signal momentum (amplitude over integration time) and therefore has the same effect on positioning error as $j(t)$. The new model, converted entirely to the discrete domain, is shown in Figure 4.

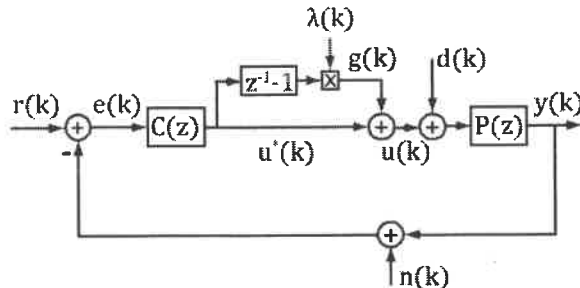


FIGURE 4. Discrete model of control output jitter with discrete modulation effect.

ANALYSIS OF CONTROL OUTPUT JITTER'S EFFECT ON POSITIONING ERROR

Two scenarios are analyzed to demonstrate the effect of control output jitter on positioning error.

First, the output regulation error resulting from jitter's interaction with the measurement noise, $n(k)$, is considered. Second, the tracking error resulting from jitter's interaction with the reference command, $r(k)$, is considered. For both analyses all other system disturbances are assumed negligible, $d(k) = 0$.

When analyzing the model from Figure 4 it is not possible to apply a direct frequency domain analysis because both $\lambda(k)$ and $n(k)$ are primarily random, thus their Fourier transforms do not exist. However, from a statistical viewpoint the autocorrelation, ϕ , of a random signal is an aperiodic sequence with a defined Fourier transform; therefore, a statistical analysis can be used to draw meaningful conclusions regarding the effect of control output jitter on positioning error.

Regulation Error Analysis with Measurement Noise

For this analysis $r(k) = 0$, although the analysis holds for any constant reference. The effect of control output jitter on regulation error must be analyzed before the measurement noise is added to the signal. Since the reference command is zero, this can be achieved by directly observing output $y(k)$. For simplicity it is assumed both $\lambda(k)$ and $n(k)$ are white. Coloured noise can be analyzed in a similar way, but fewer simplifications are possible. For white noise the following statistical conclusions can be made.

$$\phi_{nn}(k) = \sigma_n^2 \delta(k) \quad (6)$$

$$\Phi_{nn}(e^{j\omega T_0}) = \phi_{nn}(0) = \sigma_n^2 \quad (7)$$

$$\phi_{\lambda\lambda}(k) = \sigma_\lambda^2 \delta(k) \quad (8)$$

$$\Phi_{\lambda\lambda}(e^{j\omega T_0}) = \phi_{\lambda\lambda}(0) = \sigma_\lambda^2 \quad (9)$$

where $\delta(k)$ is the Dirac delta function. Using Figure 4 the system output can be expressed as

$$y(k) = y_n(k) + y_\lambda(k) \quad (10)$$

where

$$y_n(k) = -n(k) \oplus \mathcal{F}^{-1} \left[\frac{P(e^{j\omega T_0})C(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \quad (11)$$

$$y_\lambda(k) = g(k) \oplus \mathcal{F}^{-1} \left[\frac{P(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \quad (12)$$

where $\oplus$ is the convolution operation, $\mathcal{F}^{-1}$ is the inverse Fourier transform and

$$g(k) = -\lambda(k) \left(n(k) \oplus \mathcal{F}^{-1} \left[\frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \right) - a(k) \quad (13)$$

where

$$a(k) = \lambda(k) \left(g(k) \oplus \mathcal{F}^{-1} \left[\frac{C(e^{j\omega T_0})P(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \right) \quad (14)$$

Utilizing the properties for white noise, the variance of $a(k)$ can be estimated as

$$\begin{aligned} \sigma_a^2 &= \sigma_\lambda^2 \sigma_g^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{C(e^{j\omega T_0})P(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 d\omega \\ &\leq 2\sigma_\lambda^2 \sigma_g^2 \end{aligned} \quad (15)$$

Considering σ_λ is typically only a few percent, it follows that

$$\sigma_a^2 \ll \sigma_g^2 \quad (16)$$

Therefore, $a(k)$ can be neglected in Equation (13) and $g(k)$ becomes

$$g(k) \approx -\lambda(k) \left(n(k) \oplus \mathcal{F}^{-1} \left[\frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \right) \quad (17)$$

Based on Equation (10) the power spectral density of the system output can now be calculated as

$$\Phi_{yy}(e^{j\omega T_0}) = \Phi_{yy}^n(e^{j\omega T_0}) + \Phi_{yy}^\lambda(e^{j\omega T_0}) \quad (18)$$

where

$$\Phi_{yy}^n(e^{j\omega T_0}) = \sigma_n^2 \left| \frac{P(e^{j\omega T_0})C(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 \quad (19)$$

$$\begin{aligned} \Phi_{yy}^\lambda(e^{j\omega T_0}) &= \sigma_n^2 \sigma_\lambda^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 d\omega \\ &\quad \cdot \left| \frac{P(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 \end{aligned} \quad (20)$$

From Equation (20) it can be seen how control output jitter operates primarily on the high frequency controller gain to produce additional low frequency error, thus deteriorating regulation performance. The added RMS error due to jitter can be expressed as

$$\begin{aligned} \sigma_y^\lambda &= \left(\sigma_n^2 \sigma_\lambda^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 d\omega \right. \\ &\quad \cdot \left. \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{P(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 d\omega \right)^{1/2} \end{aligned} \quad (21)$$

Tracking Error Analysis with Reference Command

For this analysis $n(k) = 0$, therefore the effect of jitter on tracking error can be directly observed on the error signal, $e(k)$. Jitter is assumed to be white, so the simplifications from Equations (8) and (9) still apply. Further, the reference command considered is a single tone input,

$$r(k) = R_0 \sin(\omega_r t) \quad (22)$$

The analysis is similar to that carried out for regulation error; however, because the reference command is a deterministic sinusoidal input the autocorrelation of $g(k)$ is non-stationary,

$$\begin{aligned} \phi_{gg}(k) &= E[u_\Delta(k+m)\lambda(k+m)u_\Delta(m)\lambda(m)] \\ &= u_\Delta^2(m)\sigma_\lambda^2 \delta(k) \end{aligned} \quad (23)$$

where E is the expected value operator. From Figure 4 the tracking error due to jitter is

$$\begin{aligned} e_\lambda(k) &= -g(k) \oplus \mathcal{F}^{-1} \left[\frac{P(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right] \end{aligned} \quad (24)$$

It can be shown that $E[e_{\lambda,RMS}^2] = \overline{E[e_\lambda^2(k)]}$, thus the contribution to RMS error from jitter's interaction with the reference command is

$$\begin{aligned} E(e_{\lambda,RMS}) &= \left(\frac{R_0^2}{2} \sigma_\lambda^2 \left| \frac{C(e^{j\omega_r T_0})[e^{-j\omega_r T_0} - 1]}{1 + P(e^{j\omega_r T_0})C(e^{j\omega_r T_0})} \right|^2 \right. \\ &\quad \cdot \left. \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{P(e^{j\omega T_0})}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2 d\omega \right)^{1/2} \end{aligned} \quad (25)$$

SOLUTIONS TO MITIGATE CONTROL OUTPUT JITTER'S EFFECT ON POSITIONING ERROR

Three possibilities to reduce the additional regulation error resulting from control output jitter's interaction with measurement noise are:

- 1) Increase the controller disturbance rejection.
- 2) Add a jitter compensator.
- 3) Reduce the control output jitter.

Since the effect can be viewed as a disturbance on the control output, redesigning the controller for increased disturbance rejection will reduce the resulting error. However, design constraints and tradeoffs will impose limits on the achievable disturbance rejection. Alternately, $g(k)$ can be directly reduced by adding a jitter compensator, $C_\lambda(z)$, to the controller. The filter $\left| \frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2$ in Equation (20) has the majority of its signal energy at high frequency. $C_\lambda(z)$ attenuates this high frequency content, thus additional regulation error from control output jitter is mostly mitigated. The jitter compensator is easily realized by placing a zero at the Nyquist frequency, that is

$$C_\lambda(z) = \frac{1 + z^{-1}}{2} \quad (26)$$

This realization is ideal because it only reduces the phase margin by 5 degrees, assuming the Nyquist frequency is at least 20 times greater than the system feedback loop-transmission cross-over frequency. The final option is to reduce the control output jitter by improving or changing the control platform.

The added tracking error due to control output jitter's interaction with the reference command can only be reduced by increasing controller disturbance rejection or reducing control output jitter. The jitter compensator is not effective in this situation because the filter $\left| \frac{C(e^{j\omega T_0})[e^{-j\omega T_0} - 1]}{1 + P(e^{j\omega T_0})C(e^{j\omega T_0})} \right|^2$ in Equation (25) is evaluated at the reference command frequency, $\omega = \omega_r$, which is well below the high frequency attenuation region of the jitter compensator.

SIMULATION OF CONTROL OUTPUT JITTER'S EFFECT ON POSITIONING ERROR

To validate the control output jitter analysis a Simulink simulation is performed using the digital control block diagram from Figure 1. The continuous blocks are simulated 1000 times faster than the discrete blocks. Jitter is directly added by varying the time that the control output signal becomes available to the continuous plant block. The system parameters used for simulation are from the rotary-axial precision spindle (RAS) presented in [3]. For this system

the axial control loop is executed at 100kHz, measurement noise is 4nm RMS, and dimensionless control output jitter is 2%. Frequency responses for the plant and controller are shown in Figures 5 and 6, respectively. The controller is shown with and without the jitter compensator added.

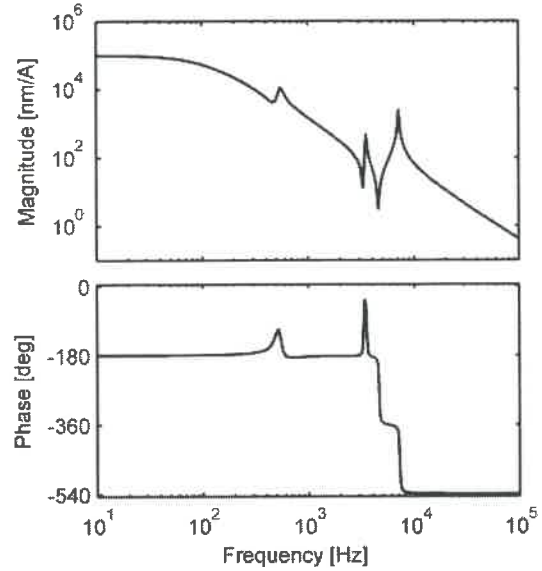


FIGURE 5. RAS plant frequency response used for simulation.

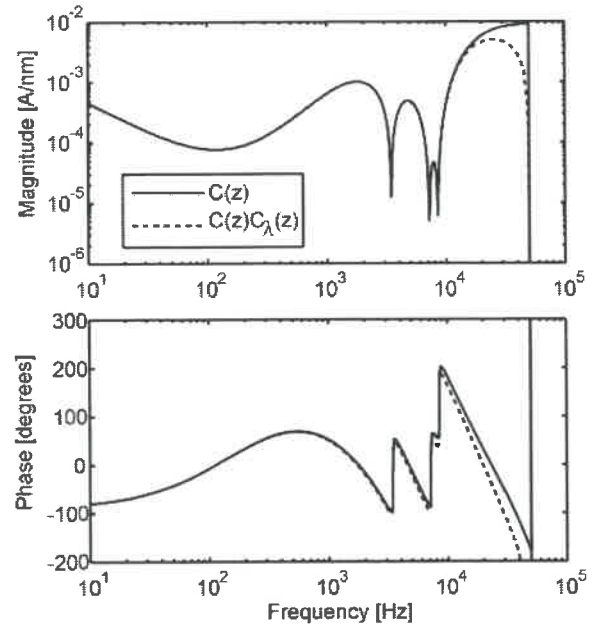


FIGURE 6. RAS controller frequency response used for simulation.

Two simulations are performed to validate the effect of jitter on both regulation error and

tracking error. The simulation parameters are listed in Table 1.

Table 1. General simulation parameters

Parameter	Value
Simulation runtime	1 s
Simulation time-step	1e-8 s
Controller sampling rate	100 kHz

Simulation Results for Regulation with Feedback Noise

For this simulation $r(k) = 0$ and the measurement noise is a random 4nm RMS white signal with zero mean. The cases considered are:

- 1) no control output jitter
- 2) 2% control output jitter
- 3) 2% control output with $C_\lambda(z)$ added

Table 2 shows the RMS regulation error on $y(k)$ for each case and compares it with that predicted from Equation (21).

Table 2. Comparison of simulated and predicted regulation error with measurement noise.

Simulation Case	Predicted $y(k)$ RMS	Simulated $y(k)$ RMS
No jitter	0.90 nm	0.89 nm
2% jitter	1.43 nm	1.42 nm
2% jitter w/ $C_\lambda(z)$	1.03 nm	1.02 nm

In Figure 7 the simulated $y(k)$ results for cases 1) and 2) are plotted to show the effect of jitter on regulation error.

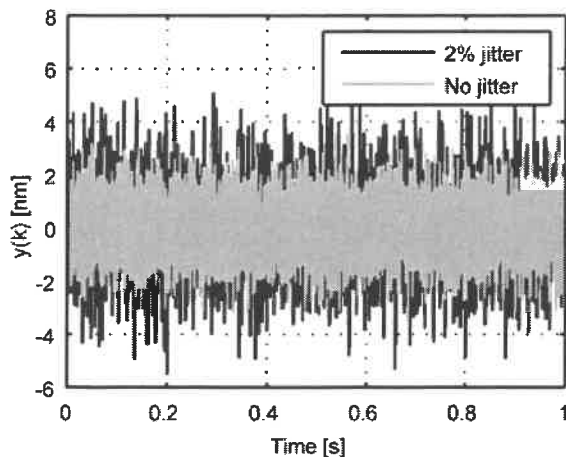


FIGURE 7. Simulated regulation error with and without control output jitter.

Note that adding 2% jitter to the system results in a 60% increase in RMS error and a 70% increase in pk-pk error. In Figure 8 the simulated and predicted power spectral densities for case 2) are compared.

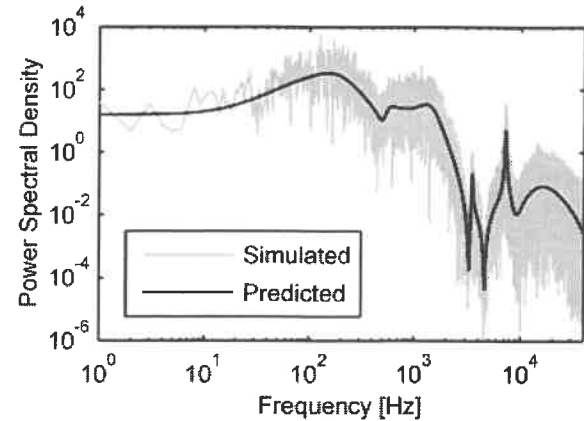


FIGURE 8. Regulation error power spectral density comparison between the predicted and simulated result for 2% jitter.

Both Table 2 and Figure 8 show that the statistical analysis accurately predicts the effect of control output jitter on regulation error when interacting with measurement noise. To provide a more thorough comparison between the simulated and predicted results the RMS regulation error on $y(k)$ is plotted versus RMS jitter in Figure 9.

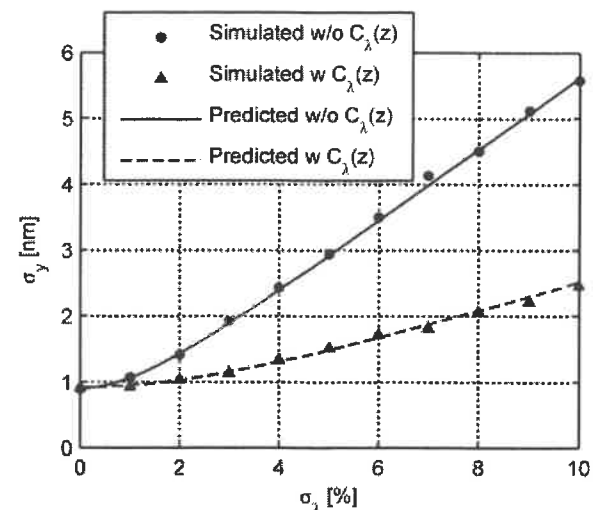


FIGURE 9. Control output jitter's effect on output regulation for 4nm RMS measurement noise.

Simulation Results for Tracking a Reference Command

For this simulation the reference command is a 5μm amplitude sinusoidal input at 1kHz and the

measurement noise is a random 4nm RMS white signal with zero mean. The jitter compensator is included for all cases to show that it is not effective for attenuating the added positioning error due to jitter's interaction with the reference command. The cases considered are:

- 1) no control output jitter
- 2) 2% control output jitter

An adaptive feed-forward cancellation (AFC) controller, discussed in [1], is incorporated so that the plant can perfectly track the reference command. Therefore all tracking error present is due to measurement noise and control output jitter. The overall predicted tracking error is calculated by summing the contribution from Equations (21) and (25). To extract the true tracking error, $e^*(k)$, from simulation $y(k)$ must be subtracted directly from $r(k)$ without adding $n(k)$ the measurement noise. Table 3 shows the true RMS tracking error for each case and compares it to the predicted value.

Table 3. Comparison of simulated and predicted tracking error with reference command.

Simulation Case	Predicted $e^*(k)$ RMS	Simulated $e^*(k)$ RMS
No jitter	0.95 nm	1.11 nm
2% jitter	3.78 nm	3.95 nm

To provide a more thorough comparison between the simulated and predicted results the true RMS tracking error is plotted versus RMS jitter in Figure 10.

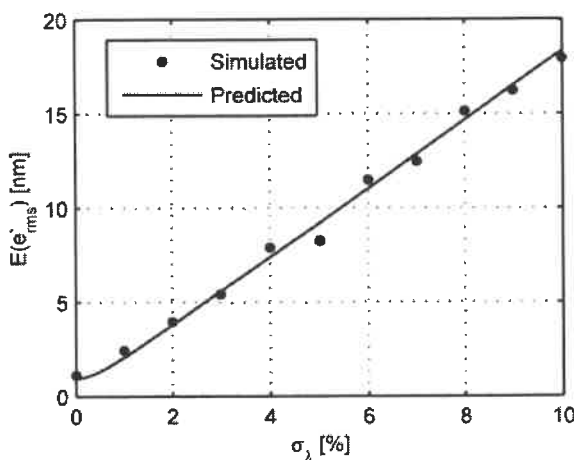


FIGURE 10. Control output jitter's effect on tracking error for a $5\mu\text{m}$, 1kHz sinusoidal reference command and 4nm RMS measurement noise.

This result shows that the statistical analysis accurately predicts the effect of control output jitter on tracking error when interacting with the reference command.

CONCLUSION

A new model is developed that includes the effect of control output jitter as a disturbance on the control output signal. This disturbance incorporates jitter as a dimensionless, zero-biased input that modulates with the derivative of the discrete control output. Based on the stochastic nature of the jitter and noise signals, a statistical analysis of the model is carried out. The result indicates that jitter interacts with both the measurement noise and the reference command to produce a low frequency disturbance, deteriorating system performance. Several solutions are suggested to reduce the effect of control output jitter on regulation and tracking error. For the case of regulation error a simple jitter compensator, consisting of a zero at the Nyquist frequency, can be used. The simulation performed using parameters from a precision magnetic bearing spindle successfully validated that the jitter model and statistical analysis accurately predict the contribution to regulation error and tracking error from control output jitter.

REFERENCES

- [1] X.-D. Lu. Electromagnetically-Driven Ultra-Fast Tool Servos for Diamond Turning. 2005; 197-224.
- [2] Lincoln B. Jitter Compensation in Digital Control Systems. Proceedings of the American Control Conference. 2002; 4: 2985- 2990.
- [3] X.-D. Lu, M.P. Paone, I. Usman, et al. Rotary-axial Spindles for Ultra-Precision Machining. CIRP Annals. 2009; 58: 1: 323-326.

TRIPOD FLOATER: AN “ACTIVE MAGNETIC BEARING” - DEMONSTRATOR

Meindert L. Norg
Norg Consulting
Canonsburg, PA, 15317 USA

INTRODUCTION

The Tripod Floater, topic of this paper, is an Active Magnetic Bearing (AMB) Demonstrator.

Goal

Goal of building this Tripod Floater is to demonstrate and increase technical capabilities and knowledge (see [4]).

Making this module completely cable-less or wireless (which is the main topic of this paper) is fueled by the same desire, more so than pursuing a specific product or application.

Active Magnetic Bearing Benefits

Some of the benefits of AMBs over traditional bearings are the ability of active vertical movements, vacuum compatibility and the fact that performance mainly depends on measurement system quality as opposed to bearing surface quality. Besides that, AMB-‘smarts’ reside in software, allowing for features like diagnostics and self-tuning capabilities.

Motivation for design choice

The specific choice for this Tripod Floater's architecture (a module with three vertical AMBs hanging from a metal plate) aims to investigate and demonstrate the (im)-possibilities of simplifying Active Magnetic Bearing designs. Most often, especially in the way AMBs have been applied in semiconductor stages over the past few years, controls architectures have been elaborate and expensive. Working towards a modular design aims to:

- Reduce price;
- Simplify design;
- Make AMB technology accessible to more engineers and a wider range of applications.

As mentioned, making the module cable-less is the main topic of this article. This required:

- An alternative controller implementation and power source;

- Implementing a wireless communication protocol;
- The use of PWM drivers instead of linear amplifiers.

Contents

This paper focuses on 4 Tripod Floater changes for making it completely cable-less. The Floater's architecture is discussed in the first section. After that, 4 changes (controller, amplifier, power source and wireless communication) will be discussed, and the section after that discusses a few ‘lessons learned’ while working on the tripod floater.

In the final section some performance results will be illuminated.

TRIPOD FLAOTER ARCHITECTURE

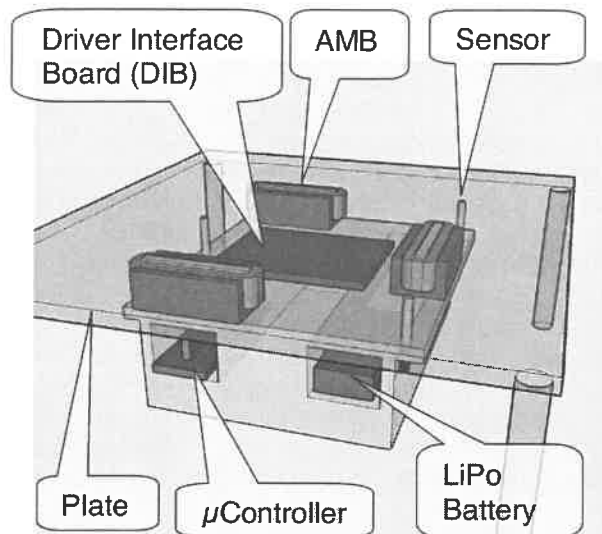


FIGURE 1. Model of cable-less Tripod Floater.

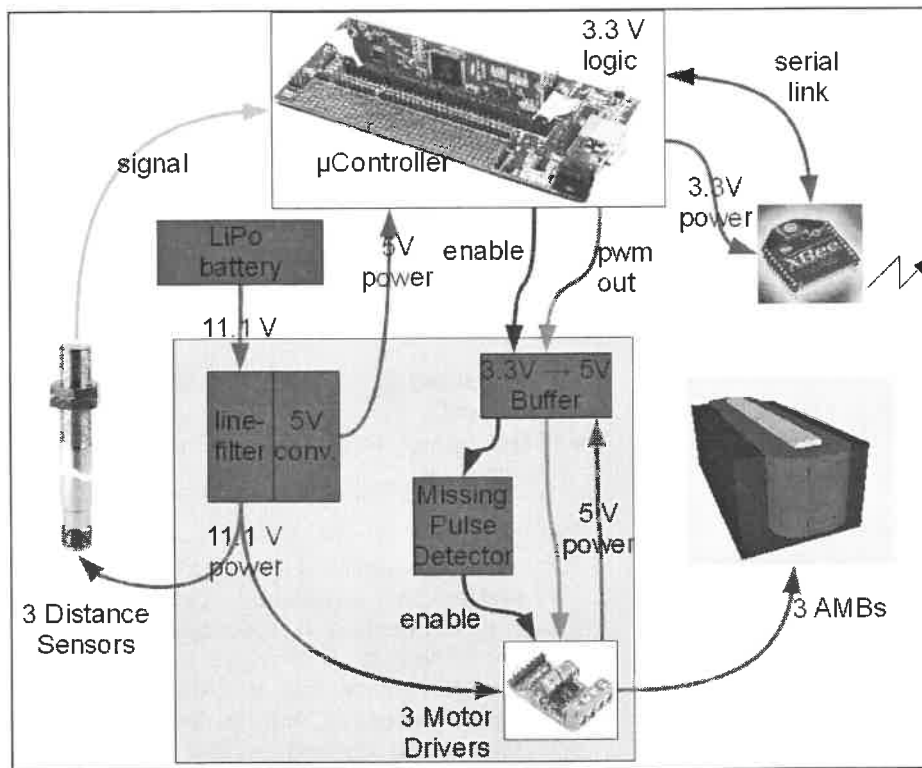


FIGURE 2. Schematics of Tripod Floater architecture.

Tripod Floater Architecture

The basic architecture of the Tripod Floater consists of an approximately 7 kg module hanging from a metal plate (see FIGURE 1) by means of 3 Electro Magnets (coils with permanent magnets, a.k.a. Active Magnetic Bearings). The distance between module and plate is measured using 3 sensors, and a sensor transformation calculates three individual gaps between the Active Magnetic Bearings (AMBs) and plate.

3 PID controllers (implemented in a μ Controller) regulate the AMB currents, maintaining a fixed distance between module and plate. The result is a constant gap between the AMBs and plate of typically 0.5 mm (for high performance applications) to 2 mm (for the Tripod Floater).

For the Tripod Floater version as described in this article, controller and battery were placed in two slots in the module's housing, and a *Driver and Interface Board* (DIB) was placed in between AMBs and sensors, as depicted in FIGURE 1 and FIGURE 6. The DIB (see also FIGURE 2) implements:

- Power filtering;
- 11.1 V $\rightarrow$ 5 V conversion;

- Sensor interfacing (power to sensors and signal to μ Controller);
- AMB driver interface;
- Additional safety and control logic.

Control Objectives

It is important to realize that the AMB-gap is actively controlled, resulting in different control objectives:

1. *Constant gap*, resulting in a friction-less multi-dimensional bearing;
2. *Vertical Reference following*, e.g. tooltip or microscope lens.
3. *Constant Current*, to minimize power consumption. In this case, the gap between the AMBs and plate would adapt to a changing load. For example, an increasing load would require an increase of magnetic force which is achievable with a decreased magnetic gap, resulting in equilibrium between load and the sum of three magnetic forces.

Bill Of Materials

In an attempt to make AMB technology more accessible an attempt was made to show it is possible to keep costs low. The sensors and AMBs themselves account for the majority of the demonstrator's costs. Electronics and mechanical components are available from www.digikey.com, www.Mouser.com and hobby stores as experimenter kits a total of about \$700.

TABLE 1. Condensed Tripod Floater BOM.

3 distance sensors	\$ 1200
3 AMBs ¹	\$ 1500
Cables	\$ 50
Motor Drivers	\$ 150
Misc. electronics components	\$ 50
2 Batteries / Charger	\$ 85
TI μ Controller Experimenters Kit	\$ 99
Zigbee Development Kit	\$ 200
Total:	\$ 3334

IMPLEMENTATION CHANGES

This section describes four major implementation changes with respect to an earlier Tripod Floater version. These changes are:

1. Move controller from external dSpace to onboard microprocessor;
2. Replace linear amplifiers with on-board PWM amplifiers;
3. Replace external power-supply with LiPo battery;
4. Replace existing Matlab-dSpace interface (for animation and tuning purposes) with wire-less Zigbee communication.

On Board Controller

For an earlier Tripod Floater version, dSpace was used to implement sensor transformations and PID controllers, allowing for faster experiments while focusing on integration.

For the current version, a Texas Instruments (www.ti.com) floating point μ Controller (TMS320F28335, see [3]) was used. A short investigation into specification in combination with experience from other projects made this an optimal choice.

A library with PIDs and 1st and 2nd order filters was built, and implemented using the following characteristics:

- 100 MHz CPU clock cycle;
- 3 PID-loops at 10 kHz;
- 12 bit ADC;
- 18 bits, 15 kHz PWM ⁱⁱ.

An Experimenter Kit docking station with separate controlCARD allows for fast and easy experimentations (see [3] and FIGURE 3).

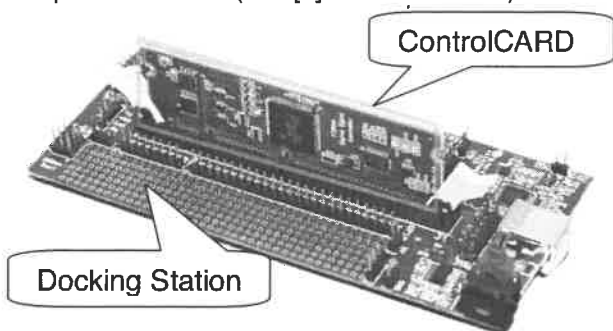


FIGURE 3. TI Experimenter Kit Docking-Station and controlCARD.

PWM Motor Drivers

To replace external linear amplifiers, three PWM motor drivers from robotics hobby parts supplier Pololu (www.pololu.com) were used.

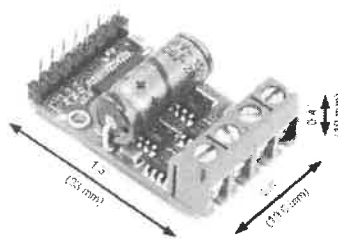


FIGURE 4. Pololu 18v15 Motor Driver.

These three drivers were ideal in size and rating to drive AMBs and easy to integrate onto the "Driver and Interface Board" (DIB) on top of the module in between all AMBs (see FIGURE 6).

Typical driver characteristics (also see [1]) are:

- 5.5 – 30 V range input voltage;
- Up to 40 kHz PWM driving frequency;
- 15 A continuous drive current (no heat sink);
- 33 x 19.6 x 10 mm.

For the controls-objective as used for this application (constant/low current), heat generation in the drivers is minimal.

LiPo Battery

A 3 cell Lithium-Polymer (LiPo) battery as often used by model-plane and -car hobbyists was used to provide power (see [2]).

Typical battery characteristics are:

- 11.1 V supply voltage;
- 2.2 Ah capacity;
- 44 A max discharge rating.



FIGURE 5. '20C' 11.1Volt Lithium Polymer battery, as often used by model airplane hobbyists.

The total power consumption of the floater is approximately 1.5 [A], most of which is used by the microcontroller and its docking station. This battery allows for up to 5 hours of experimenting time, obviously depending on the type of work.

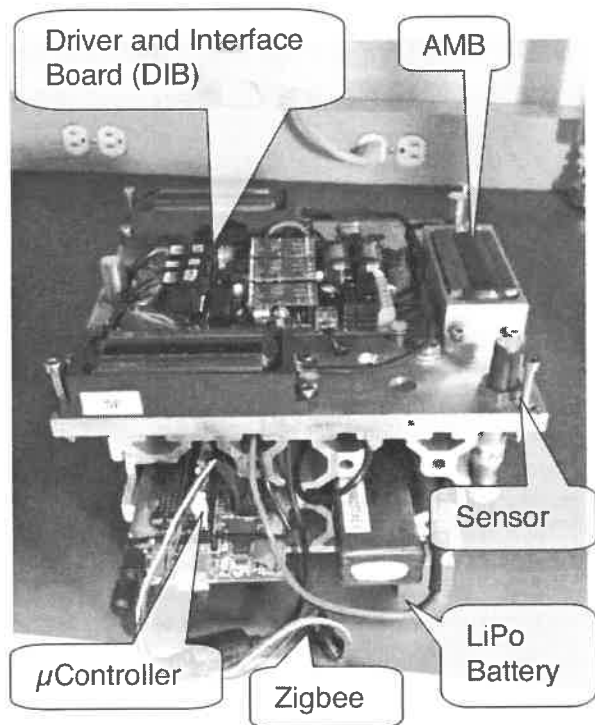


FIGURE 6. Photo of Tripod Floater standing on table.

Zigbee

To implement a means for bi-directional communication between Tripod Floater and Matlab, a Zigbee (www.digi.com) module was used (see FIGURE 7). Connected to the micro-controllers build-in serial port peripheral this provides a simple but effective communication port. The Matlab side interface is implemented using a standard Matlab COM interface.



FIGURE 7. Digi's Zigbee module.

At both sides (Matlab and microcontroller), processes continuously monitor serial port buffers for incoming data and send requests for data. This way, a baud rate of 115 kbps was achieved resulting in an animation update frequency of approximately 30 Hz.

LESSONS LEARNED

This section describes several issues that had to be addressed in order to make the Floater work:

1. AMB protection;
2. Amplifier protection;
3. Zigbee power consumption.

AMB Protection

The most basic interface between the micro-controller and the Pololu motor drivers consists of two signals:

- PWM: select PWM mode / off mode
- DIR: determines in which direction the amplifier's H-bridge are switched.

Problem

During normal operation, PWM would be ON, and DIR switches at 15 kHz, resulting in a specific desired current.

During debug sessions however, interrupting μ Controller's operation (e.g. for re-programming) might leave its outputs in a PWM-on state, and DIR in either high or low, resulting in a continuous maximum current through the AMBs. With an approximately 1 Ω resistance, this might not damage the drivers, but could overheat and damage the AMBs.

Solution

To prevent drivers from inadvertently being switched ON, an analogue 'Missing-Pulse-Detector' was implemented and connected to the μ Controller's PWM signal, which was made to pulse at a 10 Hz rate. The 'Missing-Pulse-Detector'-s function is to switch the driver's PWM input OFF when no state-change was observed for longer than 150 ms from mentioned μ Controller's output.

Amplifier Protection

Problem

During module power-up, in some cases the Motor-Driver's DIR input was set high or low before the driver was properly supplied with 5 V to power its internal logic. This would result in an unknown H-bridge state, in one case thought to have damaging the Motor-Driver beyond repairs.

Investigation

Before describing the solution to prevent this failure, it is good to know that:

- A Buffer/Driver IC is used to convert 3.3V μ Controller logic into 5V logic for the driver;
- The Motor-Drivers have a 5V output, capable of driving a few mA;

Solution

In order to prevent the Buffer/Driver to put Motor-Driver's DIR input in an undesired state a change is made in the way the Buffer/Driver is powered: Instead of powering the Buffer/Driver ICs with 5 Volts originating from the power supply, it is now powered using the Motor Driver's 5 Volt output. This way, the Motor-Driver's DIR input will not be set before the Motor-Driver is ready to receive DIR inputs.

Zigbee power consumption

Problem

During wireless communication experiments with Digi's Zigbee module, an audible hissing sound (originating from the AMB actuators) was observed each time a data-set was requested and received. Although no voltage drop was observed in μ Processor- or Motor-Driver-power supply, it was determined that the Zigbee module was temporarily drawing too much current, probably causing a temporary disturbance in controller output.

Investigation

Digi's XBee-PRO Zigbee module, powered using 3.3 Volts from the TI Docking Station, is rated to use 250 mA during data transmission.

Solution

The PRO X-Bee module was replaced with a low-power module, rated at 45 mA during transmission (see [6]).

Remaining challenges

Besides these and other issues that were fixed, several unresolved issues still require attention. As will be mentioned later, a limit-cycling effect occurs when some of the motor drivers are operated around their 0 [A] setpoint. Also, attempts to perform a proper frequency measurement have not been successful yet, for unknown reasons.

PERFORMANCE

This section describes two performance measurements.

- Stand still performance;
- Constant-Current control.

Stand still performance

As a performance indication of this modular AMB demonstrator, *FIGURE 8* shows a $10\ \mu\text{m}$ 'standstill' position error, sampled at 21 Hz using the Zigbee wireless interface.

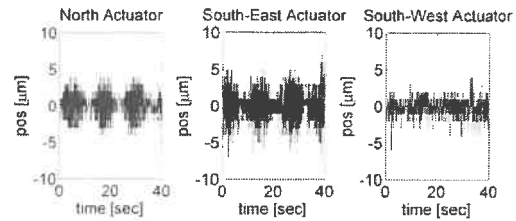


FIGURE 8. Position error during standstill. A noise level of less than $3\ \mu\text{m}_{\text{RMS}}$ is observed.

N and SE actuator 'Beating' is caused by aliasing of a higher frequent (yet to be understood) disturbance.

Constant-Current control

As mentioned earlier, one of the control objectives for an AMB module could be to carry additional pay-load at no additional current. Because of its magnetic properties, the attracting force will increase quadratically with decreasing gap.

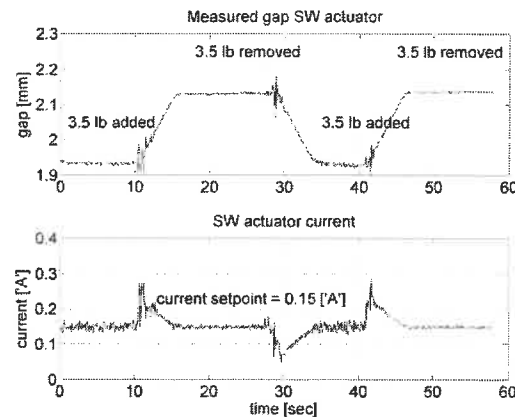


FIGURE 9. When adding and removing an additional payload a 'current controller' changes the gap (see UPPER plot) in order to maintain a current of 0.15 'Amp' (see LOWER plot).

During this experiment (see *FIGURE 9*), a mass of close to 2 kg was added and removed from the 7 kg AMB demonstrator module while floating. A current setpoint of 0.15 'Amp' ⁱⁱⁱ was chosen, to avoid the earlier mentioned limit-cycle behavior around a zero current setpoint.

After removing 2 kg of mass (at $t = 10\ \text{sec}$), the 'current controller' changes the gap setpoint from 2.13 mm to 1.93 mm in order to maintain a 0.15 'A' current. The 'current controller' simply changes the gap-setpoint with a fixed value each sample when the current is too large or too small.

CONCLUSION

In this paper, some of the work performed on the Tripod Floater, an Active Magnetic Bearing Demonstrator has been described.

Several goals were kept in mind for this project:

- Demonstrate and expand technical capabilities;
- Investigate use of PWM drivers in combination with AMB-technology;
- Making AMB technology available to a wider range of engineers and applications.

First, the general architecture has been described, and some of the work that was performed in order to make the module completely cable-less.

After that, several 'lessons learned' were listed, and finally some measurements were described.

This paper is written in the hope that technology described and modular design that is presented will assist and encourage engineers to use and experiment with 'off the shelf' components in general, and with Active Magnetic Bearings specific.

REFERENCES

- [1] Pololu Robotics and Electronics. For "High Power 18v15 Motor Driver"-specification, see:
<http://www.pololu.com/catalog/product/755>.
- [2] HobbyPartz. For battery information, see:
<http://www.hobbypartz.com/batteries.html>.
- [3] Texas Instruments. For TMS320F28335 Data Manual see:
<http://focus.ti.com/docs/prod/folders/print/tms320f28335.html>.
- [4] Norg Consulting. For more information on this and other projects, see:
www.norgConsulting.com/past_projects.html or ask Meindert@NorgConsulting.com for details.
- [5] Texas Instruments. For information on "Experimenters Kit compatible with TMS320F28335 controlCARD" see:
<http://focus.ti.com/docs/toolsw/folders/print/tmdsdock2808.html>.
- [6] Digi. For X-Bee product specifications see:
<http://www.digi.com/products/wireless/zigbee-mesh/xbee-digimesh-2-4.jsp#specs>.

ⁱ Depends strongly on tooling-costs.

ⁱⁱ 15 kHz @ 100 MHz normally would equate to a 12 bits resolution, but an 'enhanced' feature allows for a 150 ns timing resolution.

ⁱⁱⁱ The current setpoint has not been calibrated; hence its actual scaling is unknown, indicated by quotes: ['A'].

CASE STUDY FOR AN AUTOMATED LATHE ULTRAMICROTOME

Matthew Reck
Boston Engineering Corp.
Waltham, MA, USA

INTRODUCTION

A local research university sought to develop a machine that would allow them to transition from an experimental tissue sectioning process which was difficult and often unpredictable and that only skilled technicians and scientists could perform to a reliable, repeatable automated process.

APPLICATION CHALLENGES

1. Reliably slice and collect tissue samples, without damaging them, at thicknesses ranging from $1\mu\text{m}$ to 20nm .
2. Collect 10,000 or more sections from a single tissue sample over the course of several days without operator intervention.

BACKGROUND

A local university has undertaken the challenge of mapping the neural connections of the brain, with the goal of advancing the understanding of how the brain's physical connections work to create memory and thought processes. Scientists image cross-sections of brain tissue samples under high magnification using a scanning electron microscope (SEM), and then by stitching multiple images together can create a detailed 3D model of the physical sample.

After embedding tissue samples in epoxy resin, a machine equipped with a diamond knife is used to cut very thin cross-sections of the sample, typically 60 to 100 nanometers thick. This instrument is called an ultramicrotome. While several companies sell off-the-shelf ultramicrotomes that are typically used to slice a few samples at a time, the brain research project requires that thousands of sections from the same sample be sliced and collected so that the sections can be imaged and sequenced to create a 3D model of the sub-cellular structure of the tissue. Manually operated commercial ultramicrotomes are capable of only slicing a few samples at a time and require a technician with a great deal of skill, dexterity, and experience to successfully collect even 10 or 20 samples at a time. [1][2]

Researchers at the university invented a novel approach that could be used to section the tissue samples and collect them on a carrier strip. They approached Boston Engineering with their ideas and a rudimentary working prototype and asked Boston Engineering to undertake the challenge of transforming a novel idea into a robust automated lab instrument. The challenge would be to design a machine that could slice samples as thin as 20nm for extended periods of time without drift, collect the samples, and do it all with minimal operator intervention. In collaboration with the research lab, Boston Engineering has been working to design, build, and optimize a machine that will run automatically for long periods of time. Development of the cutting process and key system parameters is ongoing.

SYSTEM DESIGN

The machine developed by Boston Engineering, called the Automated Lathe Ultramicrotome, is a blend of mechanical, electrical, optical, and control sub-systems. The instrument incorporates a piezo-electric flexure stage, air bearings stages, capacitive feedback probes, and a laser interferometer. The motion controls consist of a networked servo motion controller, low noise linear amplifiers, and a sophisticated host and vision system developed using National Instruments LabVIEW.

A simplified system diagram is shown in Figure 1. The primary cutting system consists of a macro-micro linear stage stack, a goniometer, a capacitive probe feedback mechanism, a cutting knife, and a rotary stage. The cutting knife is mounted on top of the macro-micro stage stack. The sample which is to be sectioned is clamped in a chuck which is mounted to a spindle shaft oriented horizontally. The shaft is supported and rotated by a rotary air bearing stage manufactured by ABTech. TIR of the shaft is on the order of $5\mu\text{m}$. This cyclical error is repeatable and is cancelled out by means of the capacitive feedback mechanism described later.

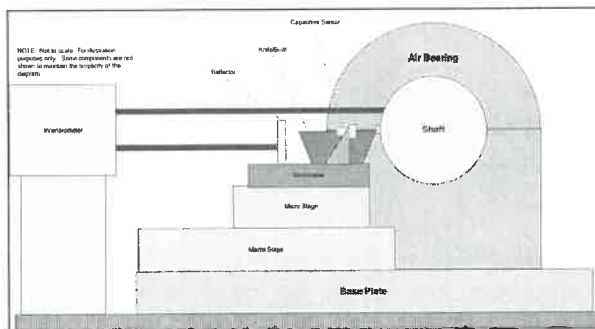


FIGURE 1. Simplified diagram of primary motion system and feedback devices.

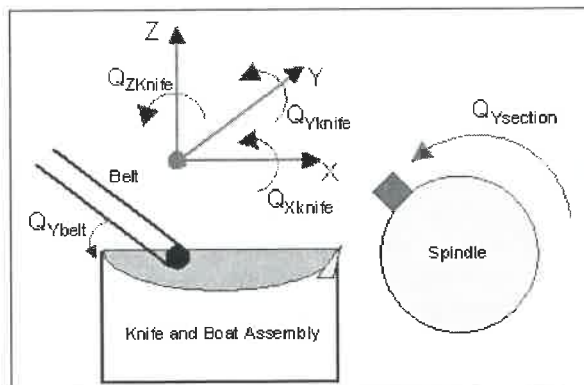


FIGURE 2. Simplified diagram of primary motion axes.

Cutting Mechanism & Stage Stack

A number of requirements led to the use of a macro-micro stage stack. The mechanical system design focused on stiffness as the best means to control the section thickness and to optimize performance, since deflections and vibrations in the system could lead to variations in cut quality. The system needed to be stiff enough to resist cutting forces so as not to deflect in either the X or Z directions. (For reference, X is the horizontal axis in which direction the knife moves towards or away from the sample. Y is the horizontal axis about which the air bearing lathe rotates. The Z axis is vertical.) The position resolution required to control the knife position relative to the sample needed to be very fine. This need for a combination of stiffness and resolution led to Physik Instrumente's line of piezo-driven flexure stages, specifically model P-733. This stage provided both X and Y motion and can be controlled via an internal feedback device or with an external device.

The primary limitation of a flexure stage is its range of travel. The system needed to be able to collect sections from a sample of over 1mm in depth, a far greater range than the $30\mu\text{m}$ travel range offered by the micro-X stage. To solve the range problem, the micro-X was mounted on the macro-X, a linear-motor air bearing stage, model ABL-1000 with a travel range of 100mm, provided by Aerotech. An air bearing stage was chosen for the typical benefits of smooth motion, cleanliness, and precision. However, the primary motivation for the use of this type of stage was the ability to turn the air supply off during section cutting. This way, the limited stiffness of the air bearing would not adversely affect the overall system stiffness during the critical cutting operation.

The cutting sequence would be thus: The rotary axis rotates the sample into the knife blade held by the micro-X stage. After each section is cut, the micro-X stage advances 20nm and the next section is cut. Once the micro-X reaches the end of its travel range, the macro-X energizes and its air supply switches on. The macro-X advances about $35\mu\text{m}$, and then the air supply is removed, effectively locking the stage in place. At the same time and while under closed loop servo control, the micro-X stage retracts its position in the opposite direction from the macro-X's advance, and in an equal amount. In this way, the position of the knife blade relative to the sample remains fixed, but the micro-X stage has been reset to its home position. The cutting operation proceeds as before until the micro-X again reaches the end of its usable range, and the reset operation is repeated.

The entire cutting mechanism, which consists of the macro-micro stage stack, the rotary air bearing, and the interferometer, were all mounted to a monolithic granite slab. Granite was chosen for its damping and stiffness characteristics and also to provide a precise reference plane to which to mount the stages. The granite slab was mounted to an optical isolation table whose main function was to shield the lathe system from environmental vibrations which may interfere with the quality of cut.

Position Control

The position of the micro-X axis relative to the sample was controlled by measuring the gap between a capacitive sensor and a precision steel shaft mounted to a rotary air bearing stage. To help account for spindle axis wobble and

noise, a pair of capacitive probes was used. Physik Instrumente's model D-510 were chosen for the capacitive gap probes. By feeding the sum of the voltage signals from the probes directly into the analog input of a motion controller, the relative position of the knife (mounted to the micro-X) and the sample (mounted to the shaft) can be maintained within the fidelity limits of the feedback device (less than 10nm).

During cutting, the motion controller attempts to hold a steady knife position based on the probe signals. If any deflections in the direction of cut are detected by the probes, a counter-motion is actuated in the micro-X stage in order to maintain a constant and consistent cut thickness. Any radial runout of the lathe axis is tracked by the capacitive probes and is compensated for by moving the micro-X axis accordingly. The bandwidth limitations of this error tracking feature have not yet been fully characterized.

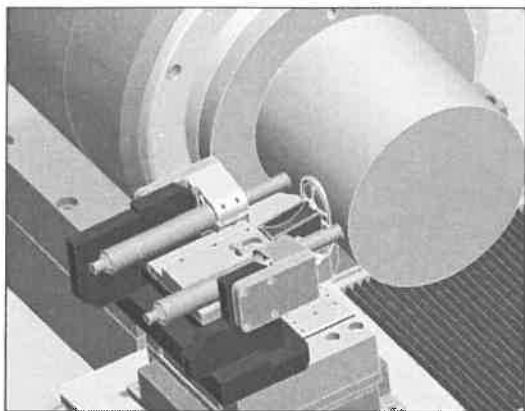


FIGURE 3. Detailed view of the capacitive gap probes. The knife holder is shown near the center of the figure.

Much like a flexure stage, capacitive probes suffer from a limited range of motion, in this case about 50 μ m of usable travel. If a 1mm section were to be sliced, the probes would only offer about 1/20th the range needed to complete an entire sample collection. To solve this problem, a sensor reset system was devised using Renishaw's RLE system.

The RLE is a fiber-launched laser interferometer capable of making differential motion measurements on the order of $\sim$ 0.4nm resolution. One plane mirror was mounted on the spindle shaft next to the sample. A second plane mirror was mounted to the rear of the

cutting knife. When the spindle shaft was aligned properly and not rotating, the relative displacement of the two optics could be reliably measured and compared to the signals provided from the capacitive probes. If the probes needed to be reset relative to the spindle shaft, the laser feedback system provided a stable reference so that the relative position of the knife to the sample could be maintained by moving the micro-X stage. Once the probes were reset, the laser and probe signals were again compared, and the cutting operation continued. The laser signal was also used to calibrate the probe signals during machine setup and testing.

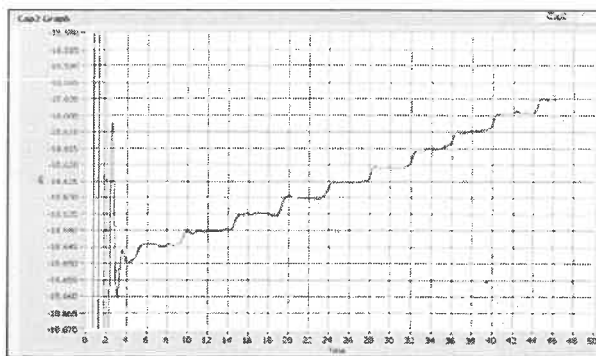


FIGURE 4. Capacitive probe data showing micro-X axis executing 5nm incremental moves.

Sample Collection

Once sliced, the tissue samples float onto the surface of a small volume of water stored in the knife fixture. The water level was maintained via a vision algorithm which measured the height of the water surface and which controlled an automated syringe pump that could add or remove water as needed. The samples were collected from the water surface by a carrier strip made of carbon-coated Kapton tape. The samples hydrostatically adhered to the tape strip, were covered with a protective clear tape, and stored on a reel. Once in this state, the samples were relatively well protected against damage or contamination and could be stored as needed.

The sample strips are later cut into short lengths, dyed, and mounted to a flat aluminum wafer. The samples are then imaged in a SEM. The physical samples can be stored indefinitely on the carrier strips and re-imaged as needed at various levels of magnification.

Control Electronics

The host control application runs on a Windows PC and communicates with two major subsystems. The first subsystem handles motion control via Aerotech's A3200 distributed network motion control system. Communications within the subsystem rely on a FireWire network, minimizing cabling and allowing high-speed data transmission back to the host PC. Position loops are closed over the network, current loops are closed locally by the servo drives.

Control of the micro-X piezo stage required a combination of components from Aerotech and Physik Instrumente to achieve the complex control via various feedback sources that was required. The piezo drive selected was PI's model E-621. This drive could control the piezo stage in standalone digital mode or in analog mode with the control signal provided externally. During simple operations such as homing, the PI controller was used in digital mode, closing the position loops of the micro-X & Y axes. In cutting mode, when the position of the micro-X needed to be maintained by way of the capacitive probes that monitored the position of the knife relative to the sample, the PI controller was switched to analog mode, and the position control signal was generated by Aerotech's NServo motion controller. This controller processed the feedback signals from the probes, generated the command signal to the piezo drive, and handled complex operations such as digital signal filtering and position calibration.

The second electronics subsystem performed data acquisition and control functions using National Instruments PXI platform. This system collected data from sensors and cameras that monitored the cutting operation in real-time. All data collected, including cutting force, stage positions, capacitive probe signals, and video, were time-synched for later analysis.

A load cell from Kistler was used to measure cutting forces and moments exerted on the micro-X stage from the interaction of the sample with the knife. The load cell measured forces on the order of 1mN in five degrees of freedom. To our knowledge, such data has never been captured before in ultramicrotomy, and future studies of the characteristics of the cutting forces could lead to a much deeper understanding of how certain materials behave when they are sliced into ultra-thin sections. [3] In the future, the data from the load cell could

also be used as an active feedback signal to further improve cutting performance.

A high-resolution CCD camera fitted to microscope optics captured high-resolution video of the sample and knife during cutting. These images can be used to determine the approximate thickness of the sample section, the sections' viability or presence of damage, and knife wear or damage. In addition to the camera, the optical system included a microscope eyepiece that the operator could use to view the work area. A multi-stage lighting system illuminated the sample and work areas from three different directions and at varying levels of intensity.

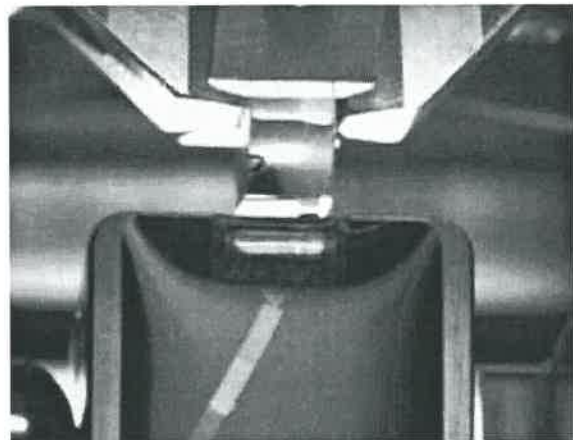


FIGURE 5. Test sample being sliced at ~30nm.

Using National Instruments DIAdem software, the force and video information can be studied post-process and analyzed concurrently to help detect relationships between the instrument data collected and the visual characteristics of the samples as they were collected.

The host control system presents a fully functional dual-screen graphical user interface to the operator that allows manipulation of the machine in both manual and automated modes. The control system is designed with numerous safety and interlock features that allow the machine to operate uninterrupted for days; weeks if necessary. All faults and alarms are logged by the system and can be sent to operators remotely via email or text message if a problem arises. By using an uninterruptible power supply and pneumatic storage tank, the system is capable of safely shutting itself down in the event of an electrical power or air supply failure without damage to the sample or the machine itself.

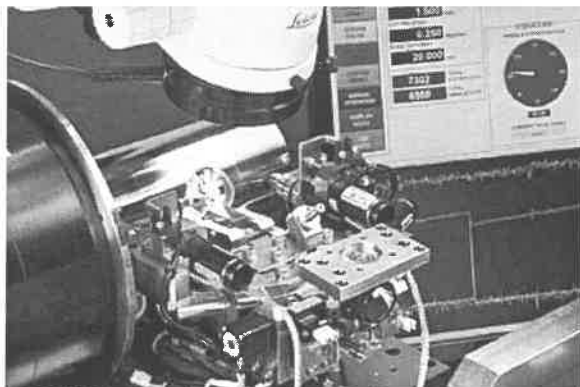


FIGURE 6. Photograph of Automated Lathe Ultramicrotome System.

ACKNOWLEDGEMENTS

The author would like to thank Ken Hayworth, Jeff Lichtman, Richard Schalek, and Bobby Kasthuri of Harvard's Lichtman Lab for their guidance and vision in the creation of this system. The author would also like to thank Martin Culpepper of the MIT Precision Compliant Systems Lab for his advice and assistance in the design of the structure and motion control system. Finally, Boston Engineering would like to credit the assistance and support provided by several key vendors and partners, including Aerotech, Physik Instrumente, Renishaw, Kistler, National Instruments, ABTech, Moscow Mills, and Newport Corporation.

REFERENCES

- [1] KJ Hayworth, N Kasthuri, R Schalek and JW Lichtman. Automating the Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microscopy and Microanalysis* (2006), 12: 86-87 Cambridge University Press, Extended abstract of a paper presented at Microscopy and Microanalysis 2006 in Chicago, Illinois, USA, July 30 – August 3, 2006
- [2] H Fernandez-Moran. Applications of a Diamond Knife for Ultrathin Sectioning of the Study of the Fine Structure of Biological Tissues and Metals. *Journal of Biophysics & Biochemistry*, 1956; Vol 2, No 4, Plates 8-15.
- [3] M Ericson and H Lindberg. Design and Potential of Instrumented Ultramicrotomy. *Polymer*, Elsevier Science Ltd, 1997; Vol 38 No 17, pp 4485-4489.

COMBINATION OF HIGH-SPEED SURFACE SCANNING AND TACTILE 3-D COORDINATE MEASUREMENT ON THE BASIS OF ADVANCED MECHATRONIC AND CONTROL CONCEPTS

Tino Hausotte¹, Eberhard Manske¹, Gerd Jäger¹, Brandon Percle¹, Nataliya Dorozhovets¹, Torsten Machleidt²

¹Faculty of Mechanical Engineering / Institute of Process Measurement and Sensor Technology

²Faculty of Computer Science and Automation / Computer Graphics Group
Ilmenau University of Technology
Ilmenau, Thuringia, Germany

INTRODUCTION

The demands on measuring and analysing micro- and nanostructures over larger ranges with increasing accuracy and precision arise from various manufacturing technologies. The measuring devices must provide multiple orders of magnitude for positioning and measurement from sub-nanometres up to hundreds of millimetres. Two tendencies can be observed. On the one hand is the realisation of surface scans in 2½ dimensions over regions as large as 350 mm x 350 mm with nanometre precision and on the other hand full three-dimensional measurements of increasingly complex structures. Conventional AFMs can carry out 2½-D surface scans with nanometre and sub-nanometre resolution but are limited to fields of 100 µm x 100 µm. The scan ranges and scan speeds must be increased. Coordinate measuring machines such as the F25 from Zeiss are specialized in 3-D micro-measurements, i.e. measurements in the nanometre range requiring very small probing elements and higher measuring resolutions. The measurements must be traceable to international standards and ultimately to the definition of the metre and they must achieve the highest possible accuracy and best measuring uncertainties.

NANOMEASURING MACHINE (NMM-1)

The Nanopositioning and Nanomeasuring Machine (NMM-1) developed at the Ilmenau University of Technology and manufactured by SIOS Meßtechnik GmbH is capable of carrying out both highly dynamic 2½-D scans and high-accuracy tactile 3-D measurements [1,3]. In this way 25 mm x 25 mm fields can be scanned with nanometre precision using various optical and tactile probe systems. 3-D point measurements and scans including freeform scans are possible in a range of 25 mm x 25 mm x 5 mm with 3-D

probe systems [1]. Positioning uncertainties of less than 10 nm are possible within this positioning and measuring volume. Until now, the resolution over the whole range has been 0.1 nm and the positioning repeatability < 0.3 nm. The basic idea behind the measurement concept is the consistent realisation of the Abbe comparator principle in all three orthogonal measurement axes [4]. This means that in all axes the specimen and the length standard are in line at all times, i.e. all three measuring axes must intersect at one measuring point. The Abbe offset can be adjusted to values below 0.1 mm by optical means. The probe system only acts as a zero indicator. The accurate reference coordinate system is defined by the corner mirror, with the interferometers measuring the position of the corner mirror and the sample directly. The interferometer measurements are used for closed-loop control to compensate the linear stage guide path errors. An additional angular control system compensates the two angular errors ϕ_x and ϕ_y using the four z-axis drives.

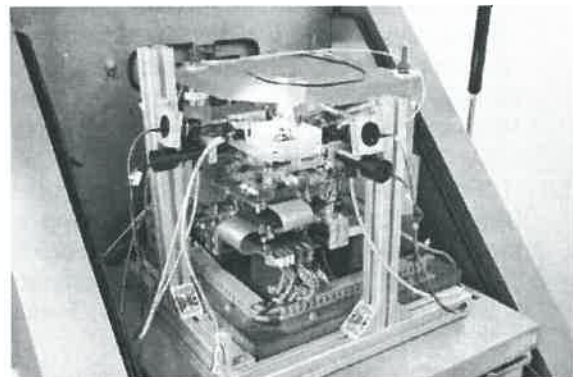


FIGURE 1. Nanopositioning and Nanomeasuring Machine (NMM-1) inside of the measurement chamber with a temperature control system and pneumatic isolators.

PROBE SYSTEMS

Small-scale surface profilers or coordinate measuring machines generally only operate using one probing system: either with an atomic force system or with a tactile micro-probe with small ruby balls. In practice, however, there are too many different requirements for sensing surfaces and for detecting structures. There are several physical principles governing the ways in which probes can be designed. Every probe exhibits specific advantages and disadvantages. Solving all these various measuring tasks together is not possible using only optical or tactile probes. It is very important to be able to select the optimum probe for a specific measurement task. Because the mechanical and electrical interface of the NMM allows the user to integrate different optical and tactile 1-D, 2-D and 3-D probe systems, many different types of measuring tasks can be carried out.

The machine can be used as a coordinate measuring machine [1], an atomic force microscope [5], an interference microscope [2], a scanning tunnelling microscope [6] and an optical and tactile surface profiler [3]. The different types of sensors are quickly and easily changeable due to their modularity. Because of the NMM's large measuring range and its very high resolution, the low attainable scan speed presents a problem. For example, an area scan with a line spacing of 100 nm at a scan speed of 100 $\mu\text{m/s}$ would require approximately two years.

SCANS WITH A FOCUS SENSOR

Conventional autofocus sensors use coil-controlled movable objective lenses. The dynamic behaviour is limited by the moving mass. A focus sensor for integration into the machine was realised on the basis of a laser hologram unit. However, in contrast to an autofocus sensor with a fixed objective lens, the focus sensor was combined with a microscope camera (see figure 2). This set-up provides a characteristic curve with a useable working range of between 2 μm and 5 μm , which can be calibrated before the measurement. The NMM's interferometric measuring system provides the capability for calibrating the probe system using the machine itself. During post-analysis the results of the focus sensor measurements are added to sign-reversed value of the z-axis interferometer. Compared to the adjusted Abbe offset of less than 0.1 mm, the tiny displacement of the probing point does not violate the Abbe

principle. The optical measurement principle provides an immediate, faster signal response than conventional autofocus sensors. Very high scanning speeds of up to 5 mm/s with uncertainties in the nanometre range can be achieved. Thus, it is possible to scan an area of 25 mm x 25 mm with 400 lines in about 30 min.

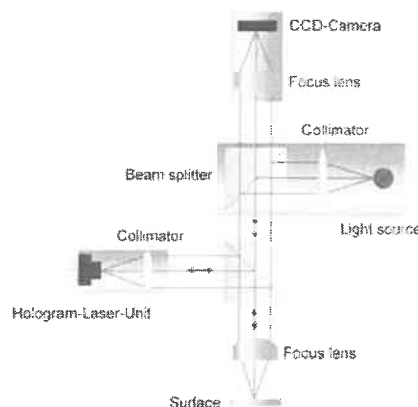


FIGURE 2. Principle of a focus sensor based on a laser hologram unit with a fixed focus lens and a microscope camera.

MEASUREMENTS WITH A WHITE LIGHT INTERFERENCE MICROSCOPE

The primary problem with the focus sensors presented in the previous section is the pointwise data acquisition, i.e., the measuring data can only be acquired sequentially point by point. Parallel data collection of 1.44 million data points (pixels) is possible through the use of white light interferometry [2]. A Mirau interference objective with a magnification of 20.0x, a working distance of 4.7 mm and a fibre-coupled halogen light source is used in this set-up (see figure 3). The measured data are acquired by a 14-bit FireWire 1394b monochromatic CCD camera at up to 30 fps.

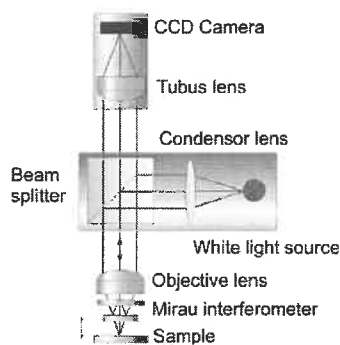


FIGURE 3. Principle of a white light interference microscope.

The system acquires a surface within a few seconds during a vertical scan of the NMM's stage in order to determine the zero optical path difference for each pixel. White light interferometry (WLI) can profit from the high positioning resolution and large measuring volume of the machine. In contrast to modern conventional white light interferometers, whose perpendicular pass-through ranges are generally limited to 100 μm , the combination of WLI with the Nanomeasuring Machine allows the principle to be applied over a height range of up to 5 mm. The measuring time can be accelerated significantly by skipping height ranges where no fringes occur.

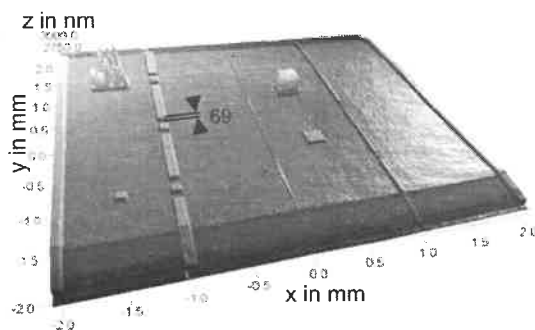


FIGURE 4. White light interference microscope measurement on a step height standard (Stitching of 64 regions).

Furthermore, the capability of exact positioning enables the user to stitch adjacent measurement areas without complex registration algorithms and independent of the topography of the sample. The field orientation and pixel scaling of the camera has to be determined relative to the machine coordinate system with capabilities of precision positioning. Samples whose lateral extensions exceed the field of view (currently about 0.8 mm x 0.6 mm) are captured by stitching height maps from several single measurements at adjacent positions. One complete measurement of the entire 25 mm x 25 mm surface can be done by stitching together approximately 2000 individual images with nanometre accuracy. This method even works if the topography of the sample is homogenous. The partial height maps can then be joined together into a single large height map independent of the topography of the sample. As an example, figure 4 shows the acquisition of the entire surface of a 69 nm step height standard (69 nm) by stitching height data maps from 64 single measurements (see figure 4). All measurements and data processing for the

resulting measuring range of 4.16 mm x 4.05 mm took about 23 minutes. The stitched height data map contains more than 21 million data points with a spacing of 0.8 μm .

2-D AND 3-D TACTILE MEASUREMENTS

Diffraction effects occur at sharp edges due to the optical principle. The optical sensors are sensitive to high contrast differences. For example, it is very difficult to measure glass scales with coated and uncoated areas. Also, different materials lead to optical phase shifts and, consequently, to faulty heights. These influences can be eliminated by using tactile sensors. Some of the tactile probe systems used only possessed one probing direction (1-D probe) and were therefore only able to be used for surface profiling (2½-D measurement). Several 3-D micro-probes were also available and were integrated into the Nanomeasuring Machine: the IBS-NPL Probe System made by the National Physical Laboratory (distributed by IBS Precision Engineering), the Gannan XP made by XPRESS Precision Engineering, a probe system made by the Institute for Microtechnology (IMT, Braunschweig University of Technology) and a probe system made at the Institute for Process Measurement and Sensor Technology (IPMS, Ilmenau University of Technology) [7].

The probing direction can be selected for the specific measuring task; it can be a fixed direction (e.g., a point measurement or a line scan) or rather a predefined change of direction (e.g., a circular scan). Additionally, a freeform scan can become possible by modifying the probing direction along a plane or cylinder wall area according to the probing force. This 3-D functionality has been successfully integrated into the NMM, bringing I++ DME measurement instruction support to the device and adding other measuring functions not covered by the specification [1]. The 3-D and freeform scanning methods especially require advanced and complex handling of force magnitude and scan motion control dependent on the force vector. These methods can be further used for surface profiling with 1-D probes for any desired probing direction (e.g. side walls of the sample). The measurement time depends on the selected measuring instruction. During point measurements the machine approaches and detaches from the sample surface at each measuring point, leading to very long measuring times. The open-loop scan method involves the probe approaching the sample once and moving

on a predefined trajectory over the sample surface. During the scan the probe system is simply deflected by surface deviations, using only the stage position as the input variable for the control system. This scan method is very fast but is limited to the probe measuring range. Another scanning method is a scan along a line or circular element using a predefined probing direction, with the probing force as the controlled variable for stage control, i.e. closed-loop control. This scan method is fast and the probe measuring range does not cause a general measurement limitation. A new scan technique developed for the NMM-1 is called 'dodge scan'. This method is essentially a combination of an open- and a closed-loop scan. The scan starts as an open-loop scan. If the deflection or probing force exceeds the set point value, the system switches to a closed-loop scan controlled using the probing force. If the system is then able to return to a position along the predefined trajectory, control is switched back to the simple open-loop control. The 'dodge scan' technique allows area scans over flat surfaces with deep holes. Freeform scan require at least 2-D probes or 3-D probes. The probing direction depends on the sample surface and is calculated from the probe system force vector and an area definition. Control is done using the probing force, allowing the probe ball to follow the surface contour. This scan method is fast and there is no measuring range limitation. The scannable contour is only limited by the diameter of the probe ball and the geometrical dimensions of the probe shaft.

ACTIVE TACTILE PROBE SYSTEMS

The probing force varies according to deflection and leads to changing interactions between the probe and the sample (see figure 5(a)). Therefore, the measuring machine must be provided with a deflection constant for each direction. Active probe systems, on the other hand, provide a constant force over the working range of the probe. There are two possible set-ups. The first is a probe with compliant deflection elements and voice coil actuators for force generation. This type of system generates an additional adjustable probing force using the voice coil actuators to keep the force constant, reacting to changes in the deflection of the spring (see figure 5(b)). The other type of active probe system is equipped with stiff deflection elements and an additional positioning system. The deflection signals are used to control the stage of the additional positioning system, which

maintains the deflection and probing force through the translation of the fastening point. This provides a constant, adjustable force within the positioning range of the stage (see figure 5(c)). For correct functionality the control system of an active 2-D or 3-D probe must be provided with the intended probing direction.

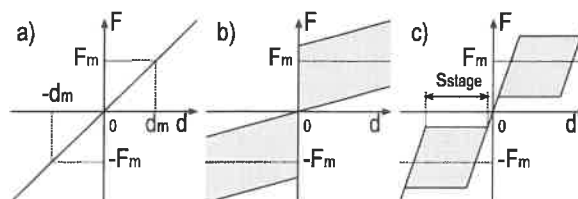


FIGURE 5. Characteristic lines of probe systems with different working principles: (a) passive probe system; (b) active probe system with a compliant deflection element and a voice coil actuator for force generation; (c) active probe system with a stiff deflection element and an additional positioning system.

An example for an active 1-D probe with an additional positioning system is the metrological AFM, equipped with a stiff and fast piezo actuator (see figure 6).

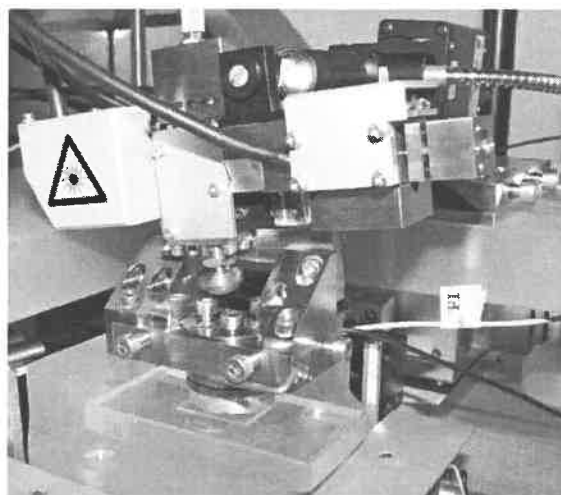


FIGURE 6. Metrological AFM in the Nanopositioning and Nanomeasuring Machine.

The normal working range with a passive probe is limited to 200 nm. In this case bending of the cantilever is controlled by the vertical motion of Nanomeasuring Machine stage and sample. Because of the high mass of about 1 kg, the control dynamics are limited. The lower mass (approx. 25 g) of the moving part of the piezoelectric drive and the cantilever allows a faster response and control of the bending. Figure 7 shows the calibration curve for the

bending (deflection measuring system) and position (interferometric measuring system) of this active probe system, along with the working range of the piezo stage. During the scans the NMM's stage executes the slow motion over the large scan length while the probe system actuator implements precise and fast force magnitude control over its 2 μm working range. The standard deviation of the bending control was reduced by factor of 10. The better force control allows much higher scan speeds.

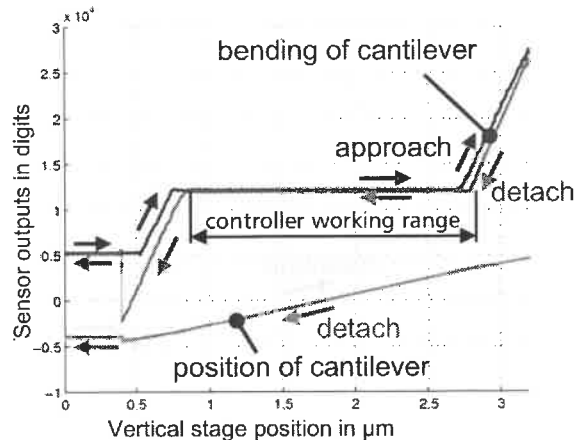


FIGURE 7. Calibration curve of the active metrological AFM.

For the realisation of active 3-D probes, the probe controller system must be provided with the probing direction in accordance with the measurement instruction. Freeform scan instructions require an especially complex force control system for the probe systems and future tests and designed are intended to work on that aspect.

CONCLUSION

This work gave an overview of the advanced mechatronic and control concepts in the NMM and the application of these concepts for the integration of probing systems. The contribution also showed the results from various measurements. Another point of emphasis here was the simultaneous use of the machine for full three-dimensional probing of the tiniest structures using highly developed 3-D microprobes.

ACKNOWLEDGMENTS

The developments presented here are supported by the German Research Foundation and the Thuringian Ministry of Education. Their support is provided within the context of

Collaborative Research Centre (SFB) 622 'Nanopositioning and Nanomeasuring Machines'. The authors wish to thank all those colleagues at the Ilmenau University of Technology who have contributed to these developments.

REFERENCES

- [1] Hausotte T, Percle B and Jäger G. Advanced three-dimensional scan methods in the nanopositioning and nanomeasuring machine. *Meas. Sci. Technol.* 20 (2009) 084004 (8pp)
- [2] Kapusi D, Machleidt T, Manske E, Franke K-H, Jahn R. White Light Interferometry Utilizing the Large Measuring Volume of a Nanopositioning and Nanomeasuring Machine. XII. International Colloquium on Surfaces, Chemnitz, 28./29.01.2008
- [3] Manske E, Hausotte T, Mastlylo R, Machleidt T, Franke K-H and Jäger G. New applications of the nanopositioning and nanomeasuring machine by using advanced tactile and non-tactile probes. *Meas. Sci. Technol.* 18 (2007) 520–527
- [4] Jäger G, Manske E and Hausotte T. Nanopositioning and measuring machine *Proc. 2nd Euspen Int. Conf.* (Turin, Italy, 27–31 May 2001)
- [5] Dorozhovets N, Hausotte T, and Jäger and Manske E. Application of the metrological scanning probe microscope for high-precision, long-range, traceable measurements. *Proc. of the SPIE* 6616 (2007) 661624 (7pp)
- [6] Weckenmann A, Hoffmann J and Schuler A. Development of a tunnelling current sensor for a long-range nano-positioning device. *Meas. Sci. Technol.* 19 (2008) 064002 (7pp)
- [7] Hofmann N, Tibrewala A, Balzer F G, Hausotte T, Manske E, Jäger G and Büttgenbach S. Characterisation of the Metrological Properties of 3-D Microprobes. *Proc. of Lamdamap* (2009) 243-252

PRECISION MACHINE CONCEPTS FOR ULTRA LARGE WORKPIECES

Bernhard Bringmann¹, Frank Brinken¹

¹StarragHeckert AG
Rorschacherberg, Switzerland

Aerospace, wind energy and nuclear equipment markets move to larger and larger workpieces with reduced tolerances. Coupled with the globalization of the supply chain for mission critical components the requirements for closer workpiece tolerances are increasing. The rigorous application of Six Sigma principles at the same time require significantly enhanced process capabilities of the machine tools applied.

With these larger workpieces the moving masses (and inertias) of a machine tool are increasing dramatically (workpiece mass, mass of machine components like pallets or columns). The stiffness of the machine cannot be increased in the same way. The in comparison low Eigen frequencies and the unbeneficial ratio of load to motor inertia result in rather bad control behavior. Still the requirements for surface finish, geometrical accuracy and process stability are ever increasing as described above.

This paper describes a novel and systematic approach for large and productive horizontal machining centers for mission critical workpieces. It is shown how theoretical concepts for achieving good dynamic behavior of such machines are put into industrial practice without compromising on the geometrical and thermal performance of the machine.

To improve the dynamic behavior of a machining center, two basic approaches are followed:

On the one hand the system can be improved to have a more beneficial correlation between excitation (e. g. machining, acceleration and friction forces) and machine reaction (vibration, displacements). This can either be done by mechanical improvements (higher stiffness, higher damping) or by improved setting of control parameters (gains etc.).

On the other hand the excitation of the machine itself can be reduced. Again this can be done by certain design realizations on the mechanical side. Furthermore it can be reduced by intelligent path planning and velocity control settings.

Different aspects in these two areas are discussed and examples for practical implementation are shown. The physical effect of such measures is explained briefly, making the connection between physical modeling and industrial best practice.

1. INTRODUCTION

In the last years a trend was established toward ever increasing part sizes in aerospace and power generation. Aircraft themselves, steam and gas turbines as well as wind energy installations grow bigger and bigger.

For StarragHeckert as a supplier of machining centers in these markets new and larger machine tools are requested. As an example there are for example the new "Big Titanium Profiler" with a pallet size of 5'000 by 2'000 mm (197" by 79") with for one single pallet alone a weight of 12 tons (see figure 1) or the new LX 351 machine for turbine blade with a length of up to 2600 mm (102") and chord length of up to 650 mm (25.5").

With this development new problems do arise concerning the control behavior of such systems. High surface finishing and geometric accuracy requirements do exist while such systems become more difficult to be controlled.



FIGURE 1. Example of a large machining center, the BTP 5000/2

The problem with growing workpieces and masses is explained in this paper. The measures for machine design and control parameter settings are explained to improve the transfer behavior between excitation and machine reaction (dynamic compliance of the machine tool). Additionally measures to reduce excitation are shown, either again by mechanical changes or with the control system - by changing path planning and velocity control settings. Finally conclusions are drawn.

CONTROL PROBLEMS WITH GROWING WORKPIECES

When workpieces are growing in dimension, like e. g. turbine blades, as a first approximation it can be said that the machine tool itself is scaled up in the same way, with pallets, bed, moving components, etc.

Therefore the masses are growing with the third power of the scaling factor. The static stiffness of a system cannot grow in a similar pace. E. g. the stiffness of a transmission belt or a ballscrew is depending on the cross sectional area, which in that approximation is growing only with the second power. The damping in the system stays more or less the same in comparison.

This of course has direct influences on the controllability of such a system.

Obvious improvements like building light weight machine structures of non-iron based materials are very difficult to implement because it would compromise the thermal stability and symmetry of the structure.

The effect of the growing masses of machine tools on the control system is shown in the example of a drive chain.

If a drive chain of a typical linear axis of a machine tool is considered, it consists of a servo motor, a coupling or transmission belt, a ball screw held by at least one bearing, a ball nut and the moving mass (e. g. column or table) attached to it (see figure 2).

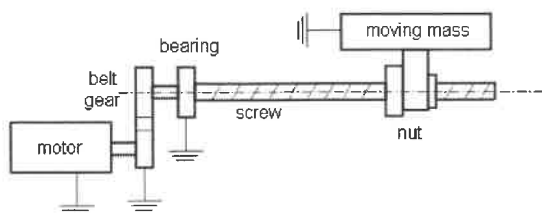


FIGURE 2. Drive chain elements of a typical ballscrew drive

The overall stiffness of this drive chain can be approximated as a serial connection of different stiffnesses, see e. g. [1]. The stiffness of the belt transmission, the axial stiffness of the bearing, the translational and the torsional stiffness of the ball screw and the axial stiffness of the nut can be added, resulting in the total static stiffness k_0 of the drive train (see figure 3) Of course there is a mix of translatory and rotary compliances which must be converted into the same unit - e.g. in equivalent rotary compliances on the motor shaft (example see [2]).

$$k_0 = \left(\frac{1}{k_{Belt,eq}} + \frac{1}{k_{Bearing,eq}} + \frac{1}{k_{Screw,eq}} + \frac{1}{k_{Nut,eq}} \right)^{-1}$$

Equally, the total inertia of such a system can be considered. Normally a distinction between the motor inertia J_{motor} (inertia of the rotor, the shaft and the first pulley of the belt transmission in the example) and the load inertia $J_{load,eq}$ is made (combined inertia of ball screw system incl. bearing and moving mass incl. nut). Of course inertias have to be defined with the same reference frame - e.g. in equivalent inertias on the motor shaft.

For the control behavior the ratio between the motor and the total inertia is important:

$$\lambda = \frac{J_{Motor}}{J_{Motor} + J_{load,eq}}$$

The first Eigen frequency can be estimated

$$\omega_0 = \sqrt{\frac{k_0}{J_{load,eq}}}$$

For this approximation, damping is neglected. In this model, the moving parts are oscillating while the rotor held in position.

The effect of such an upscaling can be seen in figure 4, starting with unit values for inertia, stiffness and damping.

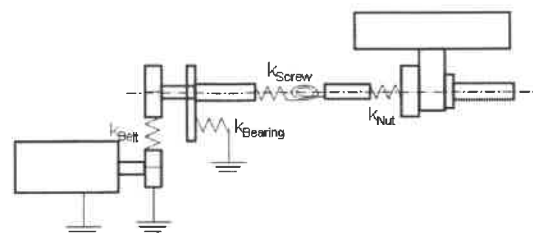


FIGURE 3. Schematic stiffnesses in a typical ballscrew drive

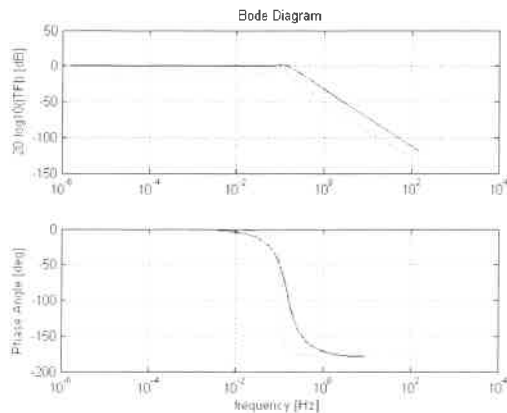


FIGURE 4. Schematic bode plot of the mechanical transfer function of a typical drive chain; original state (black) and upscaled state (grey)

The transfer function for the mechanical reaction (displacement) on a mechanical excitation (motor torque) is shown in figure 4. If a machine is scaled up by factor 2 from its original state (black line), the masses increase eightfold, the stiffnesses increase by factor 4, the damping stays more or less the same. The effect is shown in grey in the bode plot.

Of course there are higher static stiffnesses, but of course forces for reaching the same acceleration (excitation) would increase even more. The system is prone for vibration with the lower Eigen frequency and the comparably small damping.

IMPROVING THE CONTROL BEHAVIOR

Starting from such a system, ways must be found to improve the control behavior. For effectively doing this, two approaches are followed: Improving the mechanical behavior by enhancing stiffness and finding damping optimal control gains.

Improving the stiffness in the mechanical system

The goal of this approach is to bring the first Eigen frequencies of the axes to higher levels. Of course the first Eigen mode can be the overall oscillation of the drive chain as shown in the example in section 2. But it is also possible that e.g. the whole machine oscillates on the floor with a critical frequency or alternatively there could be a tilting motion of e. g. a moving column.

According to the nature of this critical Eigen mode, suitable improvements in the machine design have to be made.

In figure 1 for example it can be seen that the column has guideways not only at the bottom but also high up on the back side, impeding tilting modes.

In the example with the drive chain, an intelligent way to enhance stiffness must be found. E. g. just increasing the diameter of the ball screw has limited effects, while the overall inertia of the system is increasing as well.

At StarragHeckert, an additional thrust bearing is used with the ball screw drives.

This way, there is a parallel connection of the load to the base, resulting in a much higher stiffness k_1 :

$$k_1 = k_0 + \left(\frac{1}{k_{Screw,2}} + \frac{1}{k_{Bearing,2}} \right)^{-1}$$

If this second bearing would be just a normal thrust bearing, the ball screw could not expand when it is heating up due to friction during motion. To avoid mechanical overload of the system it must be possible to make the whole bearing movable axially when the tool is not in contact.

The mechanical implementation of StarragHeckert can be seen in figure 6.

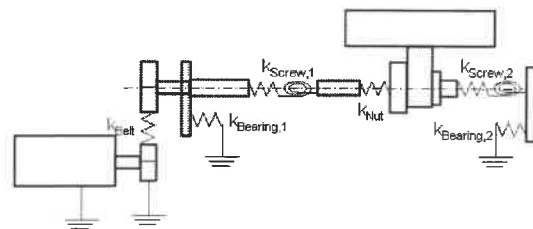


FIGURE 5. Schematic stiffnesses in a drive chain with an additional thrust bearing

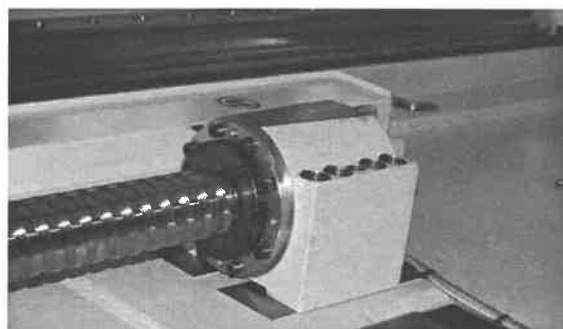


FIGURE 6. Axially unlockable thrust bearing of StarragHeckert ball screw drive

Improving damping and stiffness in the control system

For the numerical control systems in industry, the standard cascade control with current, velocity and position control loop is used. Starting from the lowest mechanical Eigen frequency, it has been shown in [2] that the transfer function of the closed velocity loop with a scaled proportional gain κ can be approximated as follows:

$$G_v = \frac{\dot{\alpha}}{\dot{\alpha}_{set}} = \frac{\frac{\kappa}{\lambda} s^2 + \omega_0^2 \kappa}{s^3 + \frac{\kappa}{\lambda} s^2 + \omega_0^2 s + \omega_0^2 \kappa}$$

With

$$\kappa = \frac{K_p}{J_{motor} + J_{load}}$$

This transfer function gives the relation between the actual velocity and the set velocity of one axis at the driving motor. The scaled proportional gain κ is responsible for the damping behavior and the setpoint response of the system. This can easily be seen by looking at the root locus plot of the transfer function

The damping brought in here is usually much bigger for machine tools than the material damping in the mechanical structure. κ must therefore be optimized for damping.

The damping D of the system can be computed by determining the angle ϕ in figure 7.

$$D = \cos(\phi)$$

With this criterion, the velocity control gain can be selected that provides optimal damping. It can be seen that the damping of the closed control loop ($D \approx 70\%$ for larger machines) is then normally much larger than the material damping in the system, even if elements with high damping values like transmission belts have been used ($D \approx 10\%$).

With higher stiffnesses in the system, higher velocity control gains can be set, leading to a faster setpoint response.

$$\kappa_{opt} \sim \omega_0$$

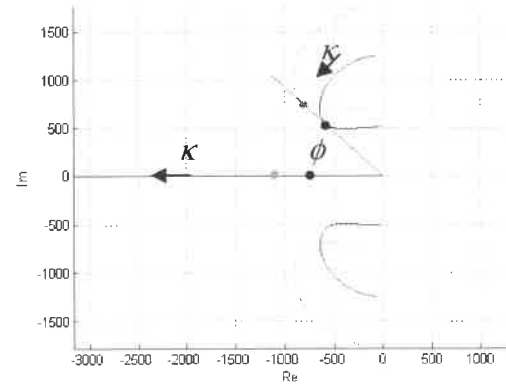


FIGURE 7. Root locus plot of the closed velocity control loop to identify damping optimal control gains; original system (black) and system with stiffness doubled (grey)

It is shown in [2] that the maximum gain in the position control loop is limited by

$$K_v \leq \frac{\kappa \omega_0}{4(\kappa + \omega_0)}$$

for direct position measurements. That means with optimal settings for the scaled velocity control loop gain κ_{opt} , the maximum proportional gain of the position control loop can be set higher with higher stiffnesses

$$K_{v,opt,max} \sim \omega_0.$$

This leads to better static load rejection e. g. of cutting forces and therefore to higher geometrical accuracy.

The different axes can be used with different values for K_v , but then velocity feed forward control has to be used.

REDUCING EXCITATION

Whatever is done to improve the mechanical system or the control parameters, systems like large machine tools always can overshoot e. g. following a step response, this way causing bad surface finishing.

After the transfer behavior between excitation and system reaction has been optimized, it must be made sure that the excitation itself stays within bounds. For finishing processes, these excitations are coming mainly from the machine itself. Mechanical and path planning measures can be taken to achieve this.

Mechanical realization for reduced excitation

When a machine tool is moving there is always a certain amount of dynamic excitation that is completely unwanted. E. g. if toothed wheels are used in a transmission, there is a multitude of pulse excitations with every tooth engagement. As a reaction the machine is oscillating in its Eigen frequencies. A similar effect is the “cogging” of linear drives.

When building a machine, such effects must be kept to a minimum. As an example other elements can be used e. g. instead of toothed wheels. Worm gear and belt transmissions do excite much less while having much higher damping rates.

Reduction of excitation of machine motion by velocity control

The biggest excitations during machining are the acceleration forces of the axes themselves. While the absolute peak magnitude of the acceleration is determining the quasi-static crosstalk effect (see [3]), the acceleration function over time is determining the dynamic excitation.

This acceleration function can either be changed by smoothing the tool path (i. e. spline interpolation) or by changing the maximum jerk values (derivative of acceleration) for the axes.

As an example if one axis should be accelerated from 0 to a constant velocity 6 m/min, and the jerk value for this axis is set to e. g. 100 m/s^3 , the acceleration function over time can be seen in figure 8 (dark curve). A maximum acceleration of 3.16 m/s^2 is reached.

In figure 9 the Fourier transformation of this signal is shown (dark line). It can be seen the principal excitation is between 0 and 30 Hz.

If the jerk value is set to e. g. 20 m/s^3 (light curve in figures 8 and 9), the maximum acceleration is reduced to 1.41 m/s^2 . It can be seen that the excitation of the low frequencies is reduced dramatically. Therefore better surface finish can be expected.

Of course reducing axes jerk values results in longer process time. To keep this time loss at a minimum, the customer can select from roughing to super finishing parameter sets on StarragHeckert machines that will among others change the jerk values to guarantee for the required surface finish while having as short as possible process times.

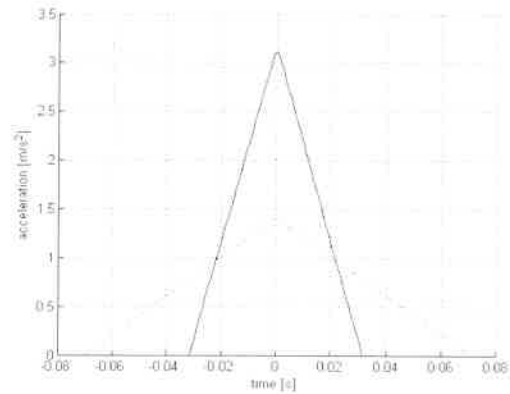


FIGURE 8. Acceleration function over time for acceleration with limited jerk values from 0 to 6 m/min; jerk limit of 100 m/s^3 (black line) resp. 20 m/s^3 (grey line)

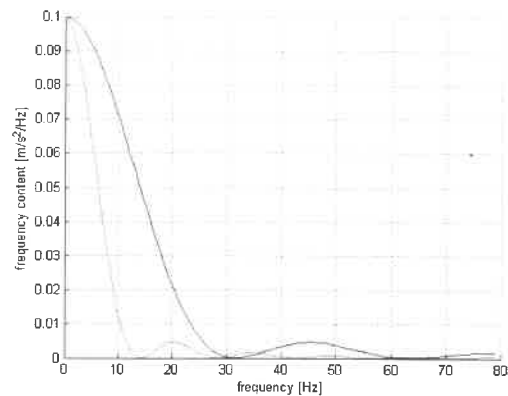


FIGURE 9. Fourier transformation of function for acceleration with limited jerk values from 0 to 6 m/min; jerk limit of 100 m/s^3 (black line) resp. 20 m/s^3 (grey line)

CONCLUSION AND OUTLOOK

With growing masses, machine tools tend to vibrate with low Eigen frequencies. To achieve high quality parts (high surface finish, high geometric accuracy) despite of that, machine and control design can be optimized – aiming at better transfer behavior as well as reduced excitation of Eigen frequencies.

Future potential lies for example in the area of control design. With state-space controllers, more damping could be introduced into the system. The velocity control could limit the jerk not to a fixed value but reduce the excitation of certain frequencies to have less excitation of the Eigen frequencies of the system.

REFERENCES

- [1] Weck M, Brecher C, Werkzeugmaschinen Mechatronische Systeme, Vorschubantriebe, Prozessdiagnose, Springer, 2006
- [2] Zirn O, Machine Tool Analysis – Modelling, Simulation and Control of Machine Tool Manipulators, Habilitation Theses ETH Zurich, 2008
- [3] Bringmann B, Maglie P, A Method for Direct Evaluation of the Dynamic 3D Path Accuracy of NC Machine Tools, Annals of the CIRP, 2009, 58/1: 339-342

HIGH-ACCURACY ATOMIC FORCE MICROSCOPE USING A SELF-SENSING PROBE

Darya Amin-Shahidi¹, Dean M. Ljubcic¹, Jerald L. Overcash², Robert J. Hocken²,
David L. Trumper¹

¹Mechanical Engineering Department
Massachusetts Institute of Technology
Cambridge, MA, USA

²Mechanical Engineering Department
University of North Carolina at Charlotte
Charlotte, NC, USA

INTRODUCTION

We present the design, control, and experimental results of a high-accuracy AFM head (HAFM). The HAFM's mechanical design was completed in the Master's thesis of Ljubcic [1]. We discuss the control and instrumentation of HAFM in this present paper, which enable sub-nanometer resolution. The HAFM is designed to be used as a high-accuracy single-axis metrology probe. To use the HAFM for metrology, it will be integrated with the Sub-Atomic Measuring Machine (SAMM), which has been designed by Holmes at the University of North Carolina at Charlotte [2]. The SAMM scans the sample in-plane over a range of 25-mm by 25-mm range with 30-nm accuracy and 1-nm repeatability. The HAFM tracks and measures the sample normal to the plane with 0.12nm RMS resolution at 100-Hz bandwidth.

To sense the surface, the HAFM uses a commercially available self-sensing Akiyama AFM probe, which is described in [3]. The self-sensing probe eliminates the need for an optical sensing mechanism and can achieve a more compact and practical design for integration with SAMM. For higher sensing bandwidth, we use the probe in a frequency-measuring mode, where the probe resonance frequency is used as feedback to control the tip-sample distance.

In this mode, a tracking controller moves the probe normal to the sample to maintain a constant self-resonance period, which corresponds to a fixed tip-sample distance. We use a piezoelectric-stack to move the probe and three capacitive displacement sensors to measure its surface-tracking motion.

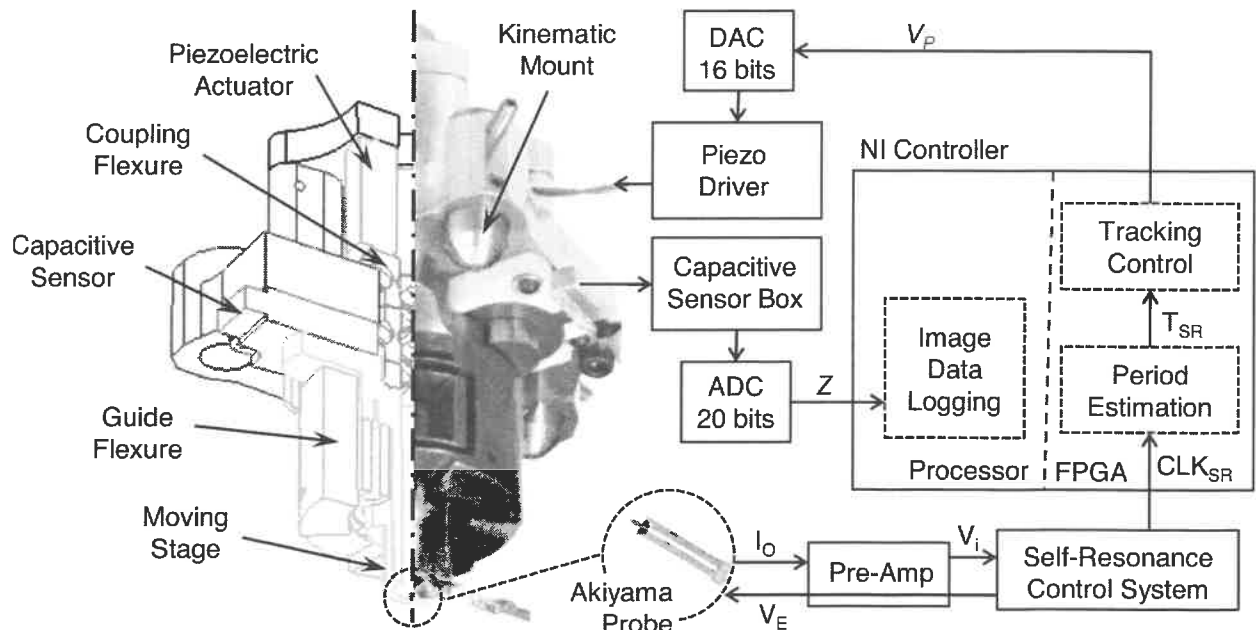


Figure 1. Diagram of HAFM system design showing the cross-sectioned head design and the electronics

SYSTEM DESIGN

The HAFM system configuration is shown in Figure 1. The HAFM sits on the SAMM's metrology frame using three ball and groove kinematic mounts. The Akiyama probe is fixed to the end of the moving stage and faces the sample surface. The HAFM *guide flexure* constrains the *moving stage* to only axial motion. A piezoelectric-stack actuator, with a range of 20 μm , drives the moving stage. A coupling flexure is used to transfer the actuator's axial motion and attenuate any non-axial error motion of the piezoelectric actuator.

The probe's output current signal, which is less than 100 nA, is amplified and converted to a voltage on a board located close to the probe. A *self-resonance control system* sets the probe in controlled amplitude self-resonance at its resonance frequency. The sinusoidal self-resonance signal is converted to a square wave with a comparator (CLK_{SR}) and is passed to the National Instruments real-time controller's FPGA board. The period (T_{SR}) of CLK_{SR} is measured and is used as feedback. The *tracking controller* commands the *piezo driver* to move the probe axially and maintain a constant self-resonance period (T_{SR}), which corresponds to a constant tip-sample gap, for a given probe excitation level.

Three ADE8810 capacitive displacement sensors, with a range of 50 μm , are symmetrically spaced around the moving stage. As HAFM tracks the sample surface, we measure its motion as the sensors' average with a resolution of 0.18 nm RMS at 1-kHz bandwidth. The controller's processor logs the sample's profile as the measured tracking motion.

AFM Probe

Two advantages of using an AMF probe for metrology are its sub nanometer resolution and localized sensing (less than 15-nm tip radius for the Akiyama probe). Many AFM systems use an optical lever sensing mechanism. To achieve a more compact design, the HAFM uses a commercially available¹ self-sensing Akiyama AFM probe, which does not need an optical-lever sensor. The Akiyama probe consists of a

¹ Akiyama probe is a product of NanoSensors™. Probe images in figures 1 and 2 are from Nanosensors, with permission.

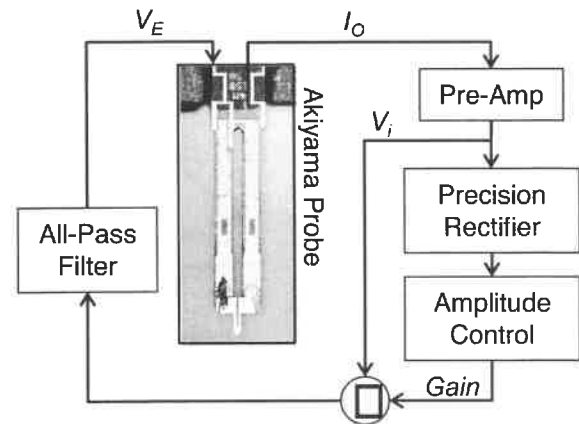


Figure 2. Self-resonance control system

quartz tuning fork and a cantilever symmetrically attached to the ends of the tuning fork's prongs. The sensing tip is located at the free end of the cantilever. As shown in Figure 2, application of harmonic voltage to the tuning fork terminals sets the tip in a tapping motion. The tip-sample interaction forces reflect back on the tuning fork's dynamics, and can be indirectly measured as the changes in the tuning fork's electrical admittance. Reference [3] describes the probe's design in more detail.

Using a tuning fork, the Akiyama probe has a high quality-factor (Q-factor). A High Q-factor improves the probe's sensitivity but reduces the decay rate of any transient. Long lasting transients can reduce the probe's bandwidth if it is used in an amplitude-measuring (AM) mode. Albrecht [4] presents a frequency-measuring (FM) AFM whose bandwidth is independent of the probe's Q-factor. In FM-AFM, the probe's resonance frequency, which shifts with the tip-sample gap, is used as feedback. The AFM controller moves the probe to maintain a reference resonance frequency, which corresponds to a fixed tip-sample gap. A review of FM and AM AFM methods is provided in [5].

We use the Akiyama probe for FM AFM in a manner similar to the design in [6]. Figure 2 shows a diagram of the self-resonance system. The current passing through the probe is converted to a voltage signal V_i . The signal V_i is scaled and phase shifted before it is applied to the probe as an excitation V_E . The signal is phase shifted to create maximum constructive feedback at the probe's resonance frequency; the amount of phase shift is empirically adjusted for each probe. This ensures that the probe self-resonates at its resonance frequency. We use an

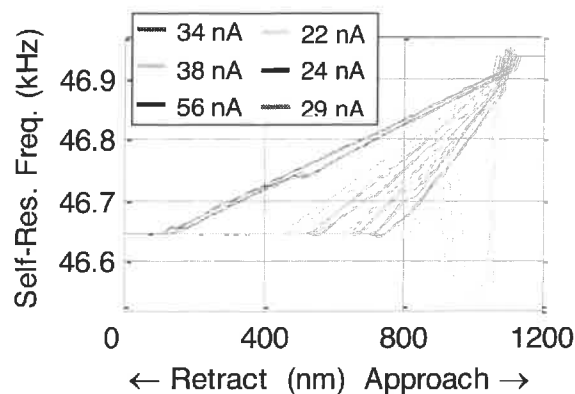


Figure 3. Sensing curves for an Akiyama probe at different oscillation amplitudes.

all-pass filter to phase-shift the signal. The loop gain is multiplicatively scaled by the amplitude controller to maintain a fixed oscillation amplitude. We use a precision rectifier to measure the amplitude of oscillations. The amplitude controller, which is implemented using op-amps, consists of an integrator in series with a lead-lag compensator.

In this configuration, the probe's self-resonance frequency, which shifts with the tip-sample gap, can be used to sense the gap. Figure 3 shows the probe sensing curves, which we captured using an Akiyama probe at different oscillation current amplitudes. We recorded the probe's self-resonance frequency over several approach-retract cycles. The retract curve is located below the approach curve and contains a discontinuity, which is likely due to probe-sample adhesion. The probe's sensitivity is inversely proportional to the oscillation amplitude. However, there is an oscillation amplitude limit below which we cannot sustain steady oscillations. This limit is 24 nA for the probe under test which corresponds to a maximum sensitivity of 0.78 Hz/nm.

SURFACE TRACKING CONTROL

The surface tracking resolution depends on the resolution of self-resonance period estimation. We use a pre-amp board, shielding on all signals, and use signal conditioning buffers to obtain a high quality self-resonance signal. The self-resonance signal is digitized using a precision comparator. A 200-MHz digital timer implemented on an FPGA measures the interval (T_{SR}) between the falling/rising edges of the digitized self-resonance signal (CLK_{SR}) with a resolution of 5-ns. A timing error of 5-ns results in an edge rate estimation error of 50-Hz for a

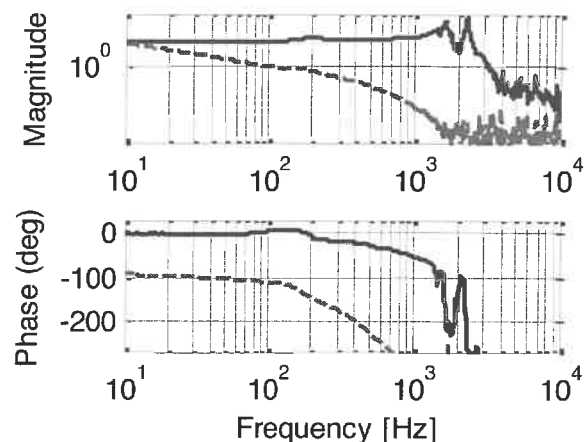


Figure 4. Bode plots of the compensated (dashed) and open loop (solid) transmissions.

probe resonating at 50 kHz. A common way to enhance the resolution is to average the period of multiple cycles, but this filtering would cause phase lag. Instead, we rely on the AFM controller and the AFM hardware to low-pass filter the period measurements. In this way, we can enhance the effective resolution and avoid additional phase lag.

Converting the self-resonance period to a frequency value requires a non-linear division operation, which would mix the frequency content. Mixing of the measurements' frequency content would prevent effective filtering of the noise. To avoid the division operation, we use the self-resonance period as feedback. In addition, the measurement and the controller update rates must have an exact integer multiplicative relation to prevent creating aliased components which would significantly degrade the system's performance. The measurement updates at the CLK_{SR} edge rate, which can change and is not exactly known. To avoid creating aliased components, we sample the

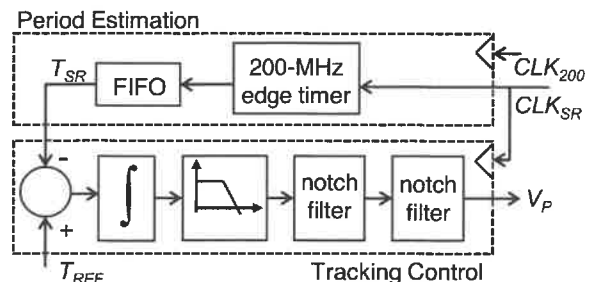


Figure 5. Diagram of the period estimation and controller blocks

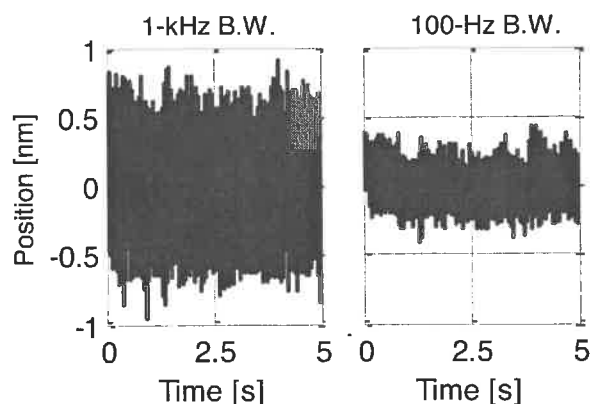


Figure 6. Tracking a stationary sample, noise is 0.24 and 0.12 nm RMS at 1000 and 100 Hz measurement bandwidths respectively.

controller at the CLK_{SR} edge rate.

Figure 4 shows the Bode plot of the open and the compensated loops from the piezo-driver command (V_P) to the probe's self-resonance period (T_{SR}). The two open-loop resonances are structural modes of HAFM and set the limit on the attainable close-loop bandwidth. Figure 5 shows the period estimation and controller blocks. The controller uses an integrator to establish a -1 slope at cross-over, a low-pass filter to attenuate the higher frequency noise, and two notch filters to mask the resonances. The unity gain loop cross-over frequency is 100 Hz with 70 degrees of phase margin.

EXPERIMENTAL RESULTS

For the tracking controller's unity cross-over frequency of 100-Hz, we tested HAFM's noise and dynamic performance at a stationary sample point. Figure 6 shows the position noise when tracking a stationary sample. Figure 7 shows the step response of the closed loop system to a step disturbance added at the piezo driver input V_P . We also successfully tested the HAFM for imaging by integrating it with a Veeco MultiMode E scanner. Figure 8 shows the images of the TGZ01 and TGX01 gratings

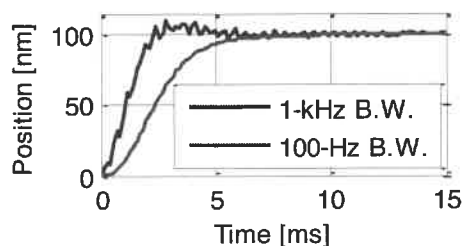


Figure 7. Tracking response to a step disturbance at piezo-driver's input (V_P).

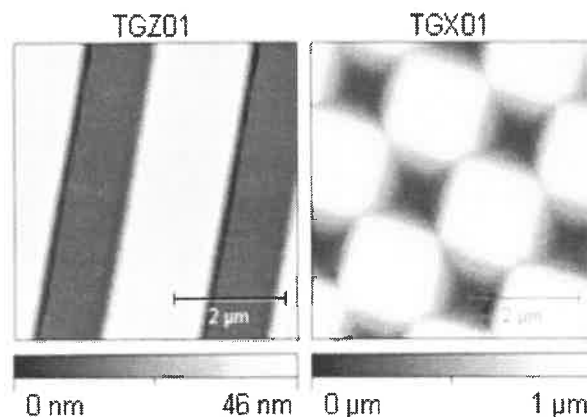


Figure 8. TGZ01 and TGX01 gratings imaged at 5 and 10 $\mu\text{m/s}$ scan speeds respectively.

captured at 5 and 10 $\mu\text{m/s}$ scan speeds respectively. The gratings are a product of MikroMasch. The TGZ01 is specified to have a channel height of 25.5 ± 1 nm, which we measured as 25.6 nm on average. The TGX01 is designed for in-plane calibration and is not precise normal to the plane. We imaged it as an example of a sample with larger features.

ACKNOWLEDGEMENT

We thank Dr. Georg Fantner and Mr. Dan Burns for giving us access to their AFM scanner system and helping us with the imaging tests. This work was supported by the National Science Foundation under contract DMI-0506898.

REFERENCES

- [1] D. L. Ljubcic, "Flexural Based High Accuracy Atomic Force Microscope," 2008.
- [2] R. J. Hocken, D. L. Trumper and C. Wang, "Dynamics and Control of the UNCC/MIT Sub-Atomic Measuring Machine," *CIRP Ann. Manuf. Technol.*, vol. 50, pp. 373-376, 2001.
- [3] T. Akiyama, U. Staufer and N. F. de Rooji, "Symmetrically Arranged Quartz Tuning Fork with Soft Cantilever for Intermittent Contact Mode Atomic Force Microscopy," *Rev. Sci. Instrum.*, vol. 74, pp. 112, January 2003.
- [4] T. R. Albrecht, P. Grutter, D. Horne and D. Rugar, "Frequency Modulation Detection Using High-Q Cantilevers for Enhanced Force Microscope Sensitivity," *J. Appl. Phys.*, vol. 69, pp. 668-673, January 15, 1991.
- [5] R. Garcia and R. Perez, "Dynamic atomic force microscopy methods," *Surf. Sci. Rep.*, vol. 47, pp. 197-301, September. 2002.
- [6] Nanosensors. (2009, 03-23). Akiyama-probe (A-probe) [technical](#) guide. pp. 1.

A USER'S GUIDE TO REPETITIVE CONTROL SYSTEMS AND THE INTERNAL MODEL PRINCIPLE

Stephen J. Ludwick and Joseph A. Profeta
Aerotech, Inc.
Pittsburgh, PA, USA

ABSTRACT

Repetitive control refers to a family of algorithms that are particularly useful when either the servo command or disturbance is largely periodic. Examples are very common in machine tools, data storage systems, and sensor testing. Any oscillatory or rotational motion generates some periodic error in both the active and ancillary motion axes. The Internal Model Principle of control theory states that algorithms designed to perfectly reject input signals must contain a model of that input, and thus a repetitive controller contains a periodic signal generator. The most general cases of repetitive control can be difficult to stabilize. If however the algorithms are limited to operate within the existing servo system bandwidth, the tuning becomes much more straightforward and the algorithms become useful tools for a servo system designer.

THE INTERNAL MODEL PRINCIPLE

The Internal Model Principle of control theory is a deceptively simple, but powerful concept. First formalized in the mid-1970's [1], it can be loosely stated as requiring that an algorithm contain a generator (or model) of any input signal if that input is to be tracked with identically zero steady state error. Figure 1 illustrates this concept with a block diagram. For there to be zero error between the commanded reference and measured signals, then the control algorithm must be able to self-generate this signal in the absence of any further input.

A familiar example of the Internal Model Principle applied in practice is through the use of an integrator (I) term in the common PID controller. Consider the case of a linear-motor-driven positioning stage modeled as a free mass with a control force applied to it. Proportional and Derivative control alone are sufficient to stabilize the system, but any constant disturbance force (due to the process, gravity, cables, and etc.) requires some error between the reference and measured positions in order for the spring-like proportional control term to generate an output. A constant disturbance is

modeled as a step input with a Laplace transform of $1/s$. Adding this term, an integrator, to the control algorithm allows the output to grow to a constant value as required to cancel the disturbance and achieve zero steady-state error.

The Internal Model Principle is very general, but specific realizations of it appear frequently in precision motion control applications. Any input (whether a command trajectory or a disturbance) that repeats with some known period can be addressed with a controller that contains a periodic signal generator. These are the repetitive controllers that will be discussed in the next section. If these inputs are frequency-limited, they can be represented as a summation of sinusoids. In that case we approach them with harmonic cancellation algorithms that apply the Internal Model Principle with a series of oscillators in the control algorithm.

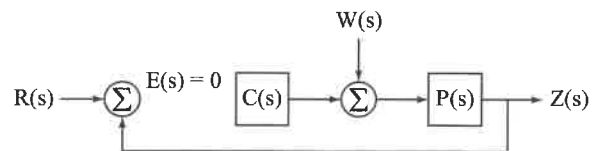


FIGURE 1. The Internal Model Principle requires that the controller contain a model of the input signals in order to be able to generate the appropriate output in the absence of any steady-state forcing error.

REPETITIVE CONTROLLERS

A periodic signal generator in the feedback control algorithm satisfies the Internal Model Principle and allows for perfect tracking of periodic commands and perfect rejection of periodic disturbances. The class of control algorithms that collectively addresses this problem is called *repetitive control*. These algorithms first appeared in the literature in the early 1980's with a paper by Inoue [2] who used the Internal Model Principle as the basis for a "controller for repetitive operation". The authors used a controller with a delay element in the feedback loop to form a periodic signal generator. In the continuous-time domain

however, the time delay element corresponded to a controller with an infinite number of marginally-stable poles. This is illustrated in Figure 2. A signal with arbitrarily-sharp transitions in the time domain requires a high-bandwidth signal generator capable of creating this high-frequency content. Stabilizing these systems is challenging since the high-frequency controller poles tend to interact with unmodeled or variable dynamics in the mechanical structure of a servomechanism which can lead to instability. A similar analysis for repetitive control algorithms in the discrete-time domain shows the same problem [3].

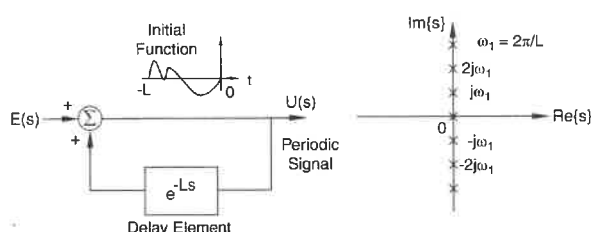


FIGURE 2. A delay element in the feedback loop of a continuous-time control algorithm satisfies the Internal Model Principle for periodic inputs, but effectively contains a high (theoretically infinite) number of oscillators to replicate an arbitrary periodic input.

The relationship between the repeating sequence in the time domain and the pole locations in the frequency domain can be understood by recalling Fourier series analyses. Any periodic signal can be equally-well represented as a summation of simple oscillating functions – namely sinusoids. Thus the repetitive controllers, when applied to linear systems, can be viewed as a series of single-frequency oscillators added to the control algorithm to cancel an input that is itself a summation of single-frequency sinusoids. This interpretation is valuable since it allows us to use a familiar tool, the Bode diagram, for determining stability margins and the steady-state response of these systems.

HARMONIC CANCELLATION

We refer to the special case of repetitive control applied to just a limited number of discrete frequencies as harmonic cancellation. These cases are very common in precision motion control applications, and include:

- Force and torque ripple
- Unbalanced payloads on rotary axes
- Cyclic command profiles
- Screw lead and gear pitch
- Link-style cable carrier systems

Notice that some disturbances can be periodic in time, while others are periodic on displacement and thus the specific frequency can vary. In the following analysis, we will assume constant-speed operation with a known frequency.

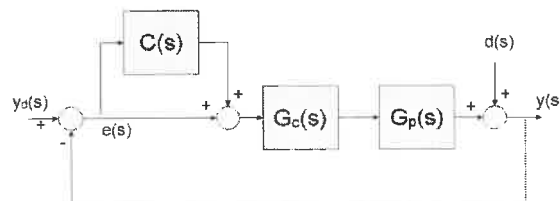


FIGURE 3. The harmonic cancellation algorithm $C(s)$ is implemented in a plug-in architecture. This keeps $G_c(s)$, the standard PID controller, unchanged and allows the harmonic cancellation algorithms to readily enabled and disabled.

We have found it useful to implement harmonic cancellation algorithms in a “plug-in” style that allows it to be easily enabled and disabled as required. The block diagram in Figure 3 represents the harmonic cancellation algorithm as $C(s)$, the standard PID controller as $G_c(s)$, and the plant (nominally a free mass) as $G_p(s)$. In keeping with the Internal Model Principle, the harmonic cancellation algorithm contains parallel oscillators, one for each frequency contained within the disturbance signal.

We can see the effect of the harmonic cancellation algorithm in the frequency domain by looking at the case of a single-frequency disturbance. Each individual oscillator in the algorithm has a continuous-time Laplace-transform representation of

$$C_n(s) = g_n \frac{s \cos \varphi_n + \omega_n \sin \varphi_n}{s^2 + \omega_n^2}. \quad (1)$$

Figure 4 shows a frequency response plot of the harmonic cancellation algorithm as the gain term sweeps from zero (disabling the oscillator) through higher values. The key point to note is that the magnitude is infinite at the oscillator frequency. Recalling our early example of the familiar integrator providing zero steady state error to constant disturbances, we can interpret

the harmonic cancellation block as simply an integrator at a non-zero frequency.

The tracking error due to disturbances is identically zero at the oscillator frequency. We can see this by referring back to the block diagram of Figure 4, and calculating that the tracking error due to a disturbance reduces to

$$\frac{e(s)}{d(s)} = \frac{-1}{1 + G_p G_c [1 + C]} \quad (2)$$

Evaluating this expression at the oscillator frequency,

$$\left. \frac{e(j\omega)}{d(j\omega)} \right|_{\omega=\omega_n} = \frac{-1}{1 + G_p G_c [1 + C(j\omega)]} = \frac{1}{\infty} = 0 \quad (3)$$

we arrive at identically zero steady-state error to disturbances at the frequency of interest. A similar analysis shows unity response with zero phase shift between the commanded and actual position profiles when the goal is to track a periodic profile.

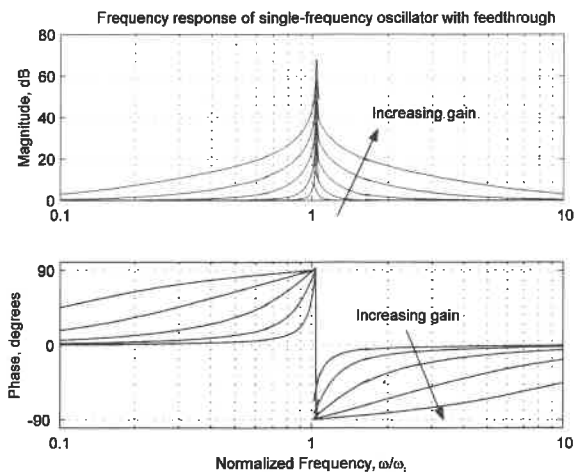


FIGURE 4. Bode plots of a single-frequency harmonic cancellation term with feedthrough show the very high (infinite) magnitude at the oscillator frequency.

In this section we have demonstrated that an oscillator in the control algorithm acts as an “integrator” term to signals at the particular oscillator frequency. Applying multiple oscillators in parallel allows the cancellation of more-complex waveforms, and approaches the general case of the full repetitive controllers. These controllers are implemented in a “plug-in”

manner that allows the standard PID control gains to remain unchanged. We use familiar frequency-domain tuning tools to determine stability margins (crossover frequency, phase margin, and gain margin) when applying harmonic cancellation algorithms. Generally however, due to the very limited frequency range that the harmonic cancellation algorithm is most active over, these systems are straightforward to tune as long as the correction frequency is well below the system crossover frequency. Figure 5 shows an experimental open-loop frequency response of an example system with the harmonic cancellation algorithm active. The dominant peaks in the loop gain below the system crossover frequency are clearly visible, and we can also see that their effect is sufficiently localized such that gain and phase at the crossover frequency are relatively unaffected.

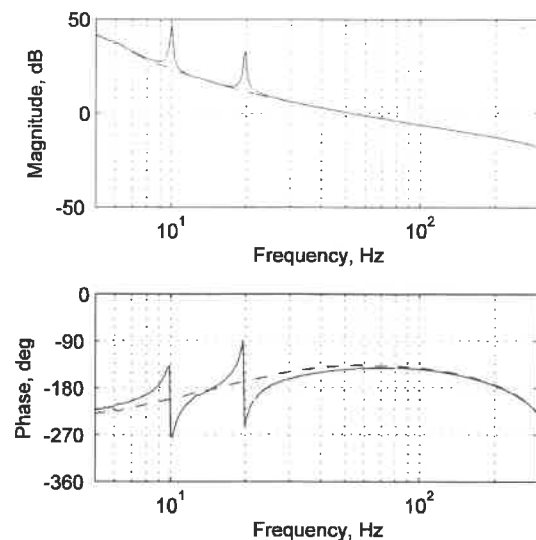


FIGURE 5. This open-loop Bode plot of a linear stage shows the strong increase in loop gain at the cancellation frequencies (10 Hz and 20 Hz) and that they have minimal influence on the response near the 55 Hz crossover frequency.

APPLICATION EXAMPLES

The overall concepts of the Internal Model Principle, repetitive control, and harmonic cancellation, once understood, are broadly applicable. One example is in the control of the read-write arm in hard disk drives [4]. The disks do not spin on a perfectly true axis, but repetitive control applied to the synchronous portion of the error motion improves the ability of the head to track the motion.

Fast tool servo mechanisms used in asymmetric turning operations benefit from the application of repetitive control as well [5][6]. When turning the surface of a toric shape (such as the mold for a contact lens that corrects for astigmatism) the cutting tool essentially returns the same point with each revolution of the spindle. This periodic toolpath can be decomposed into its Fourier series coefficients with harmonic cancellation oscillators applied to each of these.

Aerotech, Inc. designs and builds precision motion control systems, including the mechanics, drives, and control algorithms. We have found repetitive controllers to be useful enough to include them as a standard feature. The challenge for us was not in the algorithms themselves (there is a 30 year deep library of technical publications to build on) but in packaging the most-useful features into a user-friendly interface accessible to someone without an extraordinary level of training. One straightforward application was to a system with a horizontally-mounted rotary stage. The customer needed to improve their velocity stability, and we found that unbalance (once per revolution) and motor pole pitch (nine times per revolution) were the dominant terms. This is apparent in Figure 6, which shows the position error measured while the stage rotated at 60 RPM. We applied harmonic cancellation algorithms at these frequencies and reduced the root mean square tracking error from 33 arc-sec to 1.7 arc-sec, a 12x reduction.

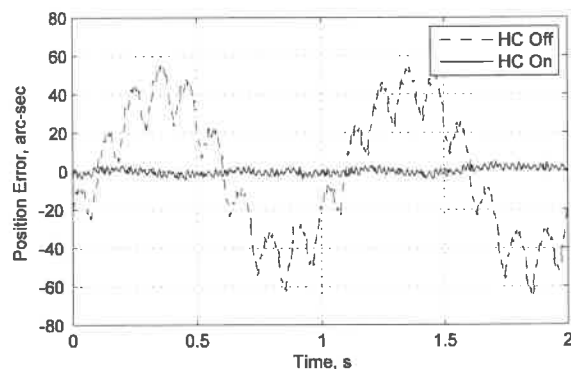


FIGURE 6. Harmonic cancellation algorithms applied to a horizontally-mounted rotary stage reduced the root mean square tracking error by approximately 12x at a 60 RPM speed. The dominant errors were the payload unbalance and torque variations at the motor pole period.

CONCLUSIONS

In this paper we presented the global concept of the Internal Model Principle of feedback control systems, and showed how this led to the development of repetitive controllers, and an even simpler case of harmonic cancellation algorithms. Periodic disturbances are commonplace in precision motion control applications, and understanding when and where these algorithms can be applied gives the control systems engineer an additional tool that is both effective, and can be analyzed with well-understood frequency-domain techniques.

REFERENCES

- [1] Francis BA and Wonham WM. The Internal Model Principle of control theory. *Automatica*. 1976;12:457-465.
- [2] Inoue T, Nakano M, and Iwai S. High accuracy control of servomechanism for repeated contouring. In *Proceedings of the 10th Annual Symposium: Incremental Motion Control Systems and Devices*. 1981;285-291.
- [3] Tomizuka M, Tsao T-C, and Chew K-K. Analysis and synthesis of discrete-time repetitive controllers. *ASME Journal of Dynamic Systems, Measurement, and Control*. 1989;111:353-358.
- [4] Chen YQ, Moore KL, Yu J, and Zhang T. Iterative learning control and repetitive control in hard disk drive industry – a tutorial. *International Journal of Adaptive Control and Signal Processing*. 2007;22(4):325-343.
- [5] Ludwick SJ, Chargin DA, Calzaretta JA, Trumper DL. Design of a rotary fast tool servo for ophthalmic lens fabrication. *Precision Engineering*. 1999;23(4):253-259.
- [6] Lu X and Trumper DL. Ultra fast tool servos for diamond turning. *CIRP Annals*. 2005;54(1):383-388.

MACHINE-TOOL TRACKING ERROR REDUCTION IN COMPLEX TRAJECTORIES THROUGH ANTICIPATORY ILC

Jon Madariaga¹, Luis G. Uriarte¹, Ismael Ruíz de Argandoña¹, José L. Azpeitia¹,
and Juan C. Rodríguez de Yurre²

¹Mechatronics Department

Fundación Tekniker IK4

Eibar, Gipuzkoa, Spain

²R & D Department

Fagor Automation S. Coop.

Mondragón, Gipuzkoa, Spain

INTRODUCTION

High precision motion control is a major issue in machine tools. Part producing with high accuracy in reduced time requires more and more challenging capabilities of these machines. However at some stage the velocity and acceleration requirements get in conflict with the dynamic capabilities of the machines and the requirements have to be limited to maintain the tracking precision.

The classical feedback control jointly with common feed-forward loops provides correct tracking of the commanded trajectories up to certain velocities/accelerations. Precisely, velocity and acceleration feed-forward loops eliminate the tracking error at constant speed and at constant acceleration respectively. The feedback loop reduces the deviations due to the effect of disturbances. However, depending on the machine dynamics trajectory commands need to be smoothed by lowering the jerk in order to limit/control the tracking error so as the desired accuracy is achieved, sacrificing overall trajectory time in the process.

The present communication proposes combining these techniques with anticipatory Iterative Learning Control (ILC) techniques to overcome that limitation when performing repetitive trajectories/parts. Moreover it can also be used to enhance the response to the not considered non-linearities and to periodic disturbances affecting the system (such as cutting forces).

LITERATURE REVIEW

ILC is based on the idea of improving trajectory tracking for repetitive processes making use of the tracking error obtained on previous iterations of the same goal trajectory.

The idea was first proposed by Arimoto *et al.* back in 1984 [1]. Since then several approaches have been reported in the literature that can be generally described as shown in equation 1, being this update law applicable to the single input, single output system denoted in equation 2 (for simplicity of notation).

$$u_{k+1}(t) = L_1(t)u_k(t) + L_2(t)e_k(t) \quad (\text{Eq. 1})$$

$$y_k(t) = G(t)u_k(t) + y_0 \quad (\text{Eq. 2})$$

where

$$e_k(t) = y_{ref}(t) - y_k(t) \quad (\text{Eq. 3})$$

being $u_k(t)$ and $y_k(t)$ represent the input and output trajectories in the run k , $G(t)$ is the operator representing the system, and y_0 the response to initial conditions, $y_{ref}(t)$ is the reference trajectory to be tracked and $L_1(t)$ and $L_2(t)$ are operators applied to the previous cycle input and tracking error trajectories, respectively and t is time.

In reference [2] an overview of several ILC approaches that fall in this general scheme can be found. The most common algorithms operate only on the error letting the operator $L_1(t)$ be the identity matrix. In this group fall the proportional ILC [3] (equation 4), the derivative ILC [1] (equation 5) and also proportional derivative controllers [4] (equation 6).

$$u_{k+1}(t) = u_k(t) + \gamma e_k(t) \quad (\text{Eq. 4})$$

$$u_{k+1}(t) = u_k(t) + \varphi \frac{de_k(t)}{dt} \quad (\text{Eq. 5})$$

$$u_{k+1}(t) = u_k(t) + \gamma e_k(t) + \varphi \frac{de_k(t)}{dt} \quad (\text{Eq. 6})$$

where γ and φ are the proportional ILC and the derivative ILC gains respectively.

However these approaches have certain drawbacks that limit their applicability in specific cases. For the case of proportional ILC controllers, their convergence has been shown to be confined to a more restricted class of systems than the derivative ILC controllers [5,6]. This feature is further explored in [7] for nonlinear continuous-time systems and update laws that consider the derivative of the error to generate new control inputs. On the other hand, derivative ILC controllers require numerical differentiation of the error signal that can lead to noise amplification if the measurement is affected by noise.

A new possibility to overcome these limitations was first proposed by Wang [8] called anticipatory ILC. He introduced the shift operator which anticipates the error of the previous iteration a certain number of samples and is then combined with a proportional operator. Since a time shift feeds the system with information from future events, the anticipatory ILC can provide derivative ILC-like robust convergence with very low noise amplification. The initial approach for nonlinear continuous-time systems with relative degree one was further developed to apply it on systems with arbitrary relative degree [9].

APPLICATION CASES

The motivation of this work is to enhance the performance of machine tools, either by improving machine precision whilst process time is kept constant or by lowering process time maintaining tracking accuracy. Two machining processes are shown as application cases in which an enhanced control technique would make a difference.

Complex trajectories such as those that can be found in machining of moulds involve major curvature changes which, when executed in short time imply significant acceleration jumps. Also direction shifts can be found that require reducing the traverse speed in order to mitigate the tracking errors produced by accelerations and decelerations that take place when changing direction.

In rigid tapping the linear displacement of a regular axis is interpolated with the movement of a highly dynamic rotating axis in order to obtain a thread of certain shape. Due to the dynamic performance difference between the two axes, the lowest axis limits the speed at which the rigid tapping can be done in order to obtain the required accuracy. When high speeds are used the different tracking errors of both axes generate a wrong thread, which can be specially noticed at the end of the hole to be tapped where high decelerations take place.

APPROACH

Anticipative ILC can help to push the capacity of the machine beyond these limitations in both complex trajectory machining and in rigid tapping.

An input update law of the following form will be used on this purpose:

$$u_{k+1}(t) = u_k(t) + \gamma e_k(t + \Delta t) \quad (\text{Eq. 7})$$

where Δt is the established anticipation time to feed data from future events.

It can be shown that provided the anticipation time is small enough the system converges towards the minimum error if the following condition is assured [9]:

$$\left\| I - \int_t^{t+\Delta t} L(\tau) C(x(\tau), \tau) B(x(\tau), \tau) d\tau \right\| < 1 \quad (\text{Eq. 8})$$

where I is the identity matrix, $L(\tau)$ is the operator to be applied to the error, $C(x(\tau), \tau)$ is the output matrix and $B(x(\tau), \tau)$ is the input matrix of the system in state space representation.

Effort has been made in order to allow an easy implementation in commercial CNCs. On this line the basic control scheme of most common CNCs (classical nested feedback loops with common feedforward loops, see Figure 1) is unaltered and the iterative learning is performed offline. The command set-point is updated every time until the desired output is obtained, maintaining the last command in the upcoming cycles. More recent algorithms make use of an additional time shift of the feedforward input [10], apart from the error time shift. However this would require feeding two data set points (one

for the regular position command and one for the feedforward command) with extensive change of the numeric control communication protocol and the drive algorithms.

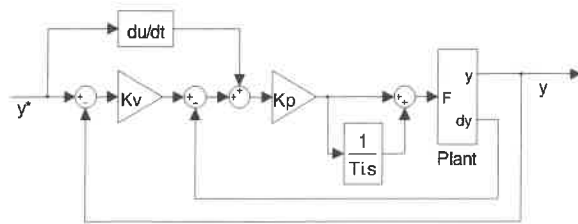


FIGURE 1. Conventional feedback and velocity feedforward control block diagram of a machine tool drive.

A characteristic anticipation time has been sought for such a system. It has been found that the settling time for 95% of the final value performs well for a wide range of dynamic systems. In order to avoid experimental work on system identification when implementing the present approach, the machine is idealized as a perfect system (see Figure 2; note that the system inertia is taken into account in gain K_p).

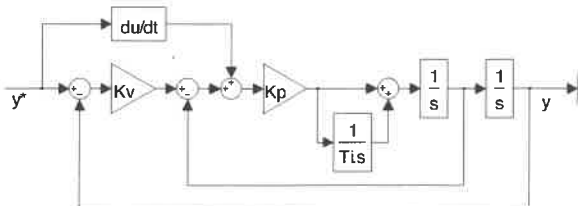


FIGURE 2. Conventional feedback and velocity feedforward control block diagram of "ideal" machine tool drive.

The ideal system's transfer function is the following:

$$\frac{y}{y^*} = \frac{K_p T_i s^2 + K_p (K_v T_i + 1)s + K_p K_v}{T_i s^3 + K_p T_i s^2 + K_p (K_v T_i + 1)s + K_p K_v} \quad (\text{Eq. 9})$$

Provided the system is already adjusted for non-ILC controlled work, the gains of the PI controller are known and the system's settling time can be obtained straightforward. The settling time of the ideal system is fairly close to the settling time of the real system if the most common adjustment criteria for machine tools are used (optimal damping and maximum stiffness – however the latter may require the use of a filter to mitigate the oscillations at the resonant frequency when ILC is applied).

Once the anticipating time is established the condition in equation 8 is applied to obtain the ILC operator. For the update law proposed in equation 7 the ILC operator is a single gain factor (γ). The proposed system is invariant in time unless the PI controller gains are changed, so the convergence condition is reduced to the following expression:

$$\|I - \gamma C B \Delta t\| < 1 \quad (\text{Eq. 10})$$

As the anticipating time is known and if the transfer function in equation 9 is formulated in state-space, the ILC gain can be isolated. To assess convergence the system is adjusted as far as possible from the limits set by the convergence condition. This value is obtained when the expression on the left-hand side in equation 10 becomes zero. So the optimal gain would be:

$$\gamma = \frac{1}{C B \Delta t} \quad (\text{Eq. 11})$$

SIMULATION RESULTS

A machine tool axis is simulated as a motor driven flexible structure with main resonant frequency at 75 Hz. An optimal damping criteria based PI controller adjustment is applied with 60 Hz bandwidth in the velocity loop and 15 Hz bandwidth in the position loop.

Two types of trajectories have been tested: a jerk limited fast positioning for rigid tapping and the complex trajectory denoted with a dashed line in Figure 3 obtained from a machined mould profile. The movement to be analyzed is the one corresponding to the Y axis of the profile (see Figure 4). The main trajectory parameters are shown in table 1.

TABLE 1. Main trajectory parameters for simulations and experimental tests.

	Fast positioning (rigid tapping)	Complex Trajectory (mould mach.)
Feed speed (m/min)	5	15
Acceleration (m/s ²)	10	5
Jerk (m/s ³)	5000	300

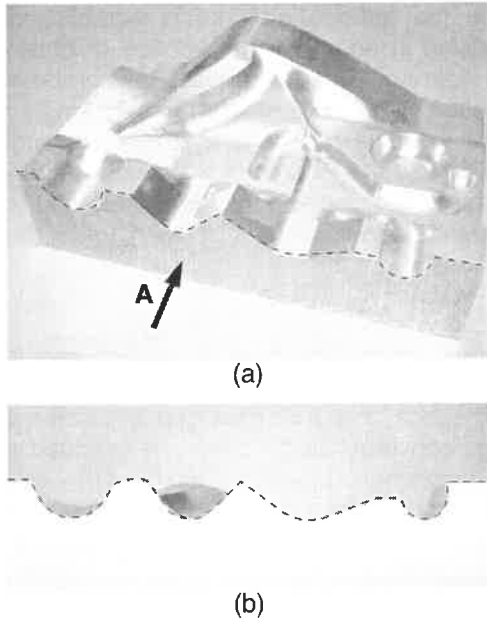


FIGURE 3. (a) Machined mould profile; (b) trajectory profile view from A.

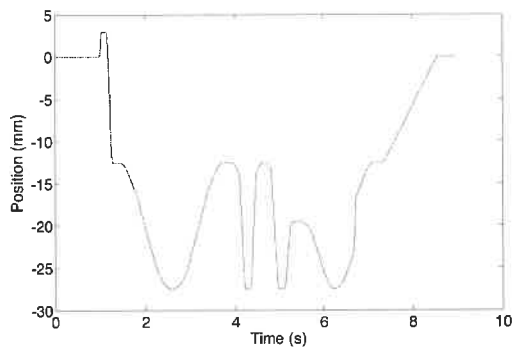


FIGURE 4. Complex trajectory for mould machining (position versus time).

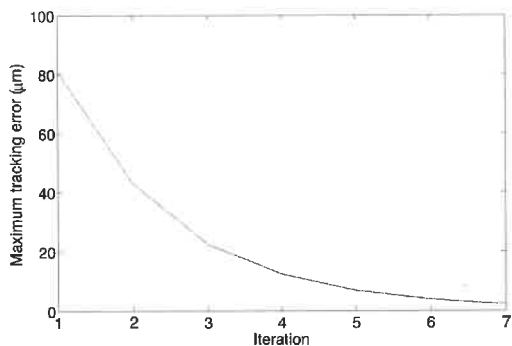


FIGURE 5. Tracking error correction in rigid tapping simulation along 7 iterations. The maximum error reduces from 80 μm to 3 μm after iteration 7.

The tracking error is improved in both cases when ILC is applied and it leads to error improvements up to 95% (see figure 5) for rigid tapping and up to 50% for the complex trajectory for mould machining (see figure 6).

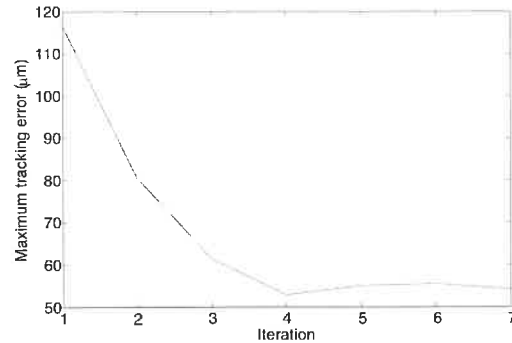


FIGURE 6. Tracking error correction in mould machining simulation along 7 iterations. The maximum error reduces from 115 μm to 55 μm after iteration 7.

EXPERIMENTAL RESULTS

The approach has been tested in a configurable test-bench specifically designed to emulate the behavior of machine-tools, which allows changing several characteristics of the drive mechanism such as friction, coupling to the motor (direct coupling or indirect transmission via pulley and belt) and stiffness of the ballscrew-nut union (see figure 7).

The system has been adjusted in order to fit with the dynamic specifications used in the simulations. The same PI controller settings employed in the simulations have been used. Both rigid tapping positioning and the complex trajectory in figure 4 have been tested with the same speed, acceleration and jerk employed in the simulations.

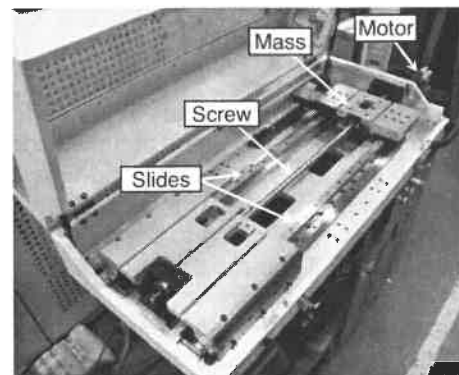


FIGURE 7. Configurable test-bench that emulates machine-tool behavior.

The obtained maximum tracking error evolution for the emulation of rigid tapping is shown in figure 8. In Figure 9 the tracking error along the whole trajectory is shown for the first and the seventh iterations. It is observed that the maximum tracking error reduces 85% for fast positioning movements similar to those in rigid tapping.

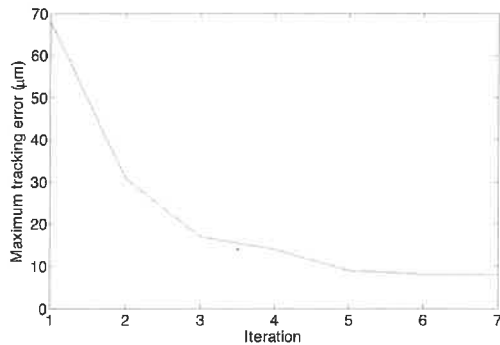


FIGURE 8. Tracking error correction in experimental emulation of rigid tapping along 7 iterations. The maximum error reduces from 68 μm to 8 μm after iteration 7.

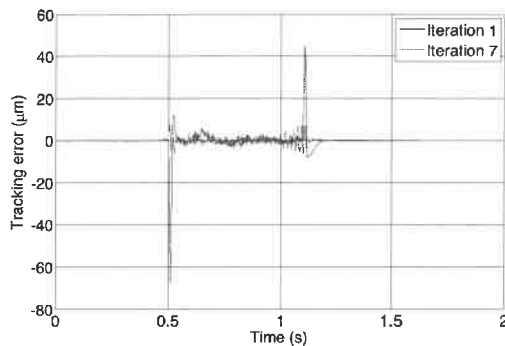


FIGURE 9. Tracking error along the whole trajectory for iteration 1 and iteration 7 in experimental emulation of rigid tapping.

Results for the complex trajectory are shown in Figures 10 and 11 (maximum tracking error evolution and tracking error along the whole trajectory respectively). The maximum tracking error reduces about 40% after 7 iterations in a trajectory similar to the ones that can be found in machining of moulds.

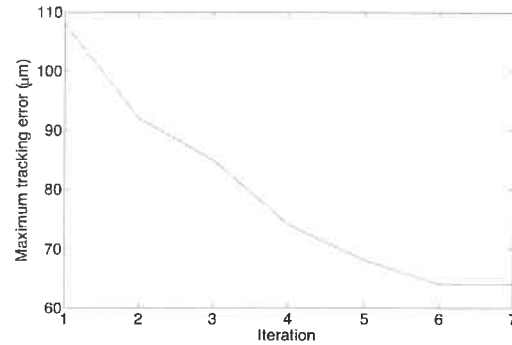


FIGURE 10. Tracking error correction in experimental emulation of mould machining along 7 iterations. The maximum error reduces from 108 μm to 64 μm after iteration 7.

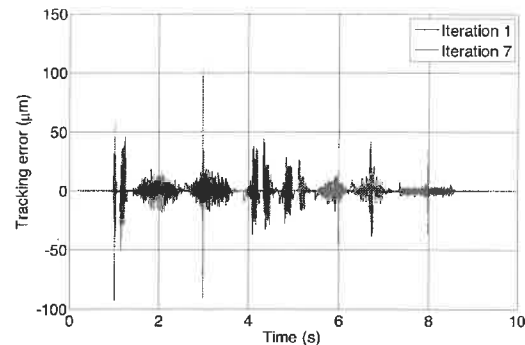


FIGURE 11. Tracking error along the whole trajectory for iteration 1 and iteration 7 in experimental emulation of mould machining.

CONCLUSIONS

Anticipatory ILC is applied to improve tracking error in machine-tools. By using the system's settling time as anticipation interval, the iterative algorithm redefines the goal trajectory after each iteration. The feeding of data from future events in a consistent way allows anticipating to upcoming tracking errors and gets closer and closer to the original trajectories, leading to accuracy improvements in the order of 80% for certain trajectories. Proper adjustment of the algorithm gains assures convergence to the minimum tracking error.

The approach has been tested in a configurable test-bench specifically designed to emulate the behavior of machine-tool drives. The experimental results show good agreement with the expected accuracy enhancement.

The applicability of the technique to real machine tools could be straightforward due to its simplicity. No specific hardware is required as it may only comprise an offline updating of the goal trajectory which is stored as position set-point once the system has converged to the desired output.

REFERENCES

- [1] Arimoto S, Kawamura S, Miyazaki F. Bettering Operation of Robots by Learning. *Journal of Robotic Systems*. 1984; 1(2): 123–140.
- [2] Ahn HS, Moore K L, Chen Y. *Iterative Learning Control*. Springer-Verlag. London: 2007.
- [3] Arimoto S. Learning Control Theory for Robotic Motion. *International Journal of Adaptive Control and Signal Processing*. 1990; 4(6): 543-564.
- [4] Kuc T Y, Nam K, Lee J S. An Iterative Learning Control of Robot Manipulators. *IEEE Transactions on Robotics and Automation*. 1991; 7(6): 835–842.
- [5] Saab S S. On the P-type Learning Control. *IEEE Transactions on Automatic Control*. 1994; 39: 2298–2302.
- [6] Sugie T, Ono T. An Iterative Learning Control Law for Dynamical Systems. *Automatica*. 1991; 27(4): 729–732.
- [7] Ahn H S, Choi C H, Kim K B. Iterative Learning Control for a Class of Nonlinear Systems. *Automatica*. 1993; 29(6): 1575–1578.
- [8] Wang D. On D-type and P-type ILC Designs and Anticipatory Approach. *International Journal of Control*. 2000; 73(10): 890–901.
- [9] Sun M, Wang D. Anticipatory Iterative Learning Control for Nonlinear Systems with Arbitrary Relative Degree. *IEEE Transactions on Automatic Control*. 2001; 46(5): 783-788.
- [10] Welz C, Srinivasan B, Bonvin D. Iterative Learning Control with Input Shift. *Proceedings of the 7th IFAC Symposium on Dynamics and Control of Process Systems*. Boston. 2004; 187-192.

COMPLEMENTARY FEEDBACK, FEEDFORWARD, AND ITERATIVE LEARNING CONTROL DESIGN FOR HIGH PRECISION REPETITIVE TRACKING PROBLEMS

Bristow, D.
(Missouri University of Science & Technology)

Many precision motion problems of practical interest involve repetitive motion. These include scanning probe microscopy, wafer stepping, robotic welding, and rapid prototyping machines. One effective precision motion control technique for tracking repetitive motions is Iterative Learning Control (ILC). ILC is an adaptive feedforward control technique that uses the tracking errors from previous iterations to construct the feedforward control signal. The learning process converges after anywhere from a few to tens of iterations, depending on the algorithm. Typically, ILC (when combined with the system's existing control system) eliminates nearly all repetitive errors, reducing the RMS tracking error by one or two orders of magnitude.

One well-known drawback of ILC is that it cannot compensate for non-repetitive disturbances. In fact, non-repetitive disturbances and noise are amplified by ILC because the learning algorithm will attempt to compensate for them on subsequent iterations (when they have changed or no longer exist). Thus, there is a tradeoff in ILC between repetitive error reduction and non-repetitive error amplification. Generally, it is not possible to completely isolate the repetitive error from non-repetitive error via filtering. Feedback control, however, can effectively compensate for disturbances (repetitive and non-repetitive), but is bandwidth-limited and noise-amplifying. Feedforward control is high-bandwidth and noise insensitive, but is limited by model accuracy and only applies to measurable disturbances (or the reference). Therefore, the control designer has three tools (feedback control, feedforward control, and ILC), each with its own set of (overlapping) capabilities and drawbacks.

This work decouples the design problem to create a tractable procedure for complementary design of the feedback, feedforward, and ILC to minimize overall error. The approach combines recent results in frequency domain stochastic ILC analysis with classic feedback and feedforward loop-shaping ideas.

NON-RIGID BODY FEED-FORWARD FOR HIGH PRECISION STAGES

**Michiel J. Vervoordeldonk, Mark Baggen
ASML, De Run 6665, 5504 DT Veldhoven , The Netherlands**

This work discusses a feedforward concept for high precision stages that takes into account the non-rigid body nature of the stages by compensating for the mechanical deformations that occur during the acceleration phase of a motion profile.

In industrial high precision motion systems such as positioning devices in semiconductor industry, the use of both a feedforward controller and a feedback controller as part of the controller architecture is common practice. The feedback controller provides disturbance reduction, while the feedforward controller is used to improve the tracking performance. The feedforward algorithms are usually based on a rigid body model of the stage, as is for instance the case with the conventional acceleration feedforward.

Tracking performance degrades with increasing set-point acceleration levels. With the continuously increasing demand for more accurate and faster motion systems, it becomes important to take into account the non-rigid body character of a stage in the feedforward path.

The presentation discusses a novel non-rigid body feed forward concept based on the modal description of the internal deformation of the stage. It will be shown how existing dynamic models are used in a novel way. The new feedforward results in adjusted positioning signals, rather than in additional controller forces, as is the case with conventional feedforwards. As the feedforward output is proportional with the inverse of the stiffness terms of the flexible modes, the concept is called compliance compensation – or cc for short. The method is compared with the so called jerk derivative feedforward algorithm (a non-rigid body feedforward concept presented at the ASPE spring topical meeting in 2004) where feedforward forces proportional with the fourth derivative of the set-point are generated.

The differences as well as the strengths of both concepts will be addressed. Both methods have their specific advantages. As it turns out the feedforward performance can be further enhanced by combining cc and jerk derivative feedforward, resulting in superior tracking performance. Simulations of industrial relevance show that the proposed concept significantly contributes to the realization of faster and more accurate motion systems.

THE MYTHICAL ACTIVE VIBRATION ISOLATION

J. van Eijk^{1,2}, D.A.H. Laro³, J. Eisinger³, W.W.J. Aarden³, C.J.M van den Berg³

1. Technical University Delft,
Delft, The Netherlands

2. MICEbv,
Eindhoven, The Netherlands

3. MI-Partners,
Eindhoven, The Netherlands

SUMMARY

Precision Systems have always been threatened by disturbances originating from vibrations. Active Vibration Isolation Systems have promised significant improvements of precision. Many different concepts have been introduced. Broad application however has not yet occurred. Actual benefits appear not sufficient to accept the added cost. Many aspects contribute to this situation. Three of these aspects will be presented at Euspen's Annual Meeting in June 2010. These aspects are, The following elements will be discussed;

1. Specification of floor induced vibrations
2. Noise Generation in active systems
3. Vibrations induced by driving forces

A short introduction to these elements will be given and special attention will be given to the next three categories of complicating effects;

4. Non-linearity in suspension system and vendor's presentation of performance graphs
5. Sensor sensitivity to cross-talk
6. Sensor sensitivity to gravitation

NON LINEARITY IN ACTIVE ISOLATION SYSTEMS.

The floor specifications based upon the standards proposed by Bolt, Beranek and Newman, BBN, have been used for many years now. They specified different curves with different levels of vibration suited for different environments. The curves are indicated as BBN-A to BBN-E.

In the design of a precision system the integral influence of vibrations can be predicted using the Dynamic Error Budgeting procedure, Lit 2. In this approach the floor vibrations are described using the Power Spectral Density functions, PSD. Prediction of the resulting vibration on the suspended mass can be done with the use of the

Transfer Function from floor-vibration to payload vibration. Figure 1 shows an example of such a transfer curve that indicates a huge performance over a broad range.

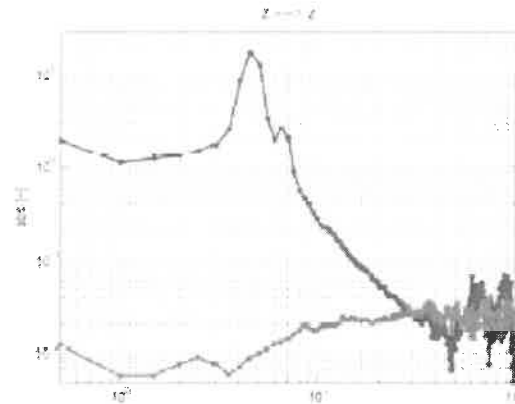


FIGURE 1. Example of transfer function for passive (blue) and active isolation system. (source Lit.1)

The basic approach in Dynamic Error Budgeting is founded on theories for linear systems and stochastic signals. When extremely good isolation is desired levels of vibration will be small and it is important to assess the required linearity. For lower frequencies, say upto 1 Hz, the level of acceleration of the floor is around $10e-5 \text{ m/s}^2$. When the active system is reducing the acceleration with a factor 100 the resulting acceleration will be less than $10e-7 \text{ m/s}^2$.

When we assume a suspended mass of 1000 kg this means that the forces acting on the payload are smaller than 0.1 mN! The suspension is creating forces of 10,000 N to support the mass. No friction forces, pre-sliding stiffness or hysteresis effects can be accepted. Extremely small values of disturbances must be eliminated. In practical realizations this is extremely difficult to achieve.

CROSS TALK FROM ACTUATOR TO SENSOR

To achieve good performance at lower frequencies some form of Geophone is used to measure the pay-load velocities. In the Geophone a small magnet is suspended in springs. Around the magnet coils are placed and the relative velocity from magnet to coil will induce a back-EMF. The measured voltage on the coils will be used as the velocity feedback signal. Due to the stiffness of the flexures used for the suspension the sensor will measure accurately for frequencies above the resonance frequency, typically about 4 Hz. Below that frequency the relative velocity between magnet and coil due to the input velocity decreases dramatically. At 0.4 Hz the relative velocity is only 1/100 times the input velocity and as a result the measured voltage will be very small.

In the active vibration isolation actuators are used to counteract payload vibrations. In many cases these actuators are of electro-magnetic nature. Lorentz type actuators have favorable characteristics. When generating forces these actuators will influence the magnetic fields in their surroundings. As the coil is, in general, sensitive to fluctuations in magnetic fields the driving currents in the actuators may cause voltages to appear on the coil. Although these voltages are small in the normal operating region they may become significant at low frequencies.

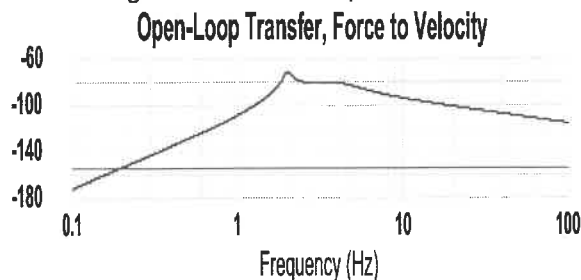


FIGURE 1. The Open Loop Transfer from Force to Geophone output and an estimated cross-talk from motor current to Geophone Voltage.

During the design of the feedback controller stability issues may arise at the 0 dB cross-over frequencies. From the transfer function of driving force to measured velocity shown in figure 2 it is clear that there will be two such frequencies. The high frequency one will be around 40 Hz, while the low frequency cross-over will be around 0.2 Hz. At this low frequency the Geophone sensitivity will be very small, thus the influence from actuator currents at these frequencies may significantly influence the phase of the transfer function. Such unexpected phase differences

have led to stability problems at low frequencies which made the system fail to start up.

SENSOR SENSITIVITY TO GRAVITY

In the Geophone the velocity of the magnet relative to the coils is measured. When the sensor rotates the gravity force acting on the magnet will also cause a displacement of the magnet. Small rotations of a vertically mounted sensor can generally be neglected. For horizontal sensors this effect is however quite significant. The voltage measured at the coil will be proportional to the angular velocity of the sensor. As a result the different control directions will become strongly coupled and may introduce non-minimal phase behavior.

To handle this issue different alternatives are available. Most solutions depend on the proper combinations of all sensor information and thus reconstructing the appropriate information. Such an approach is acceptable and interesting for a research environment. For more practical cases the development at MI-Partners has resulted in a pragmatic solution that is easily implemented and does allow for straightforward control solutions.

CONCLUSION

Active Vibration Isolation systems hold promises to improve the positioning performance of Precision Systems. The extent of the improvement promised is generally over-estimated. The mysteries and myths surrounding this matter have seriously hampered the broader industrial acceptance.

REFERENCES

- [1] G.W van der Poel, J. van Dijk, H.J.M.R. Soemers and J.B. Jonker Selection of motion sensors for active vibration isolation of hard-mounted payloads, Proceedings MOVIC 2008
- [2] Jabben L. Mechatronic Design of a magnetically suspended rotating platform, 2007, PhD Thesis TUDelft
- [3] TNO and Mecal, 2009, The Netherlands, <http://www.mikrocentrum.nl/FilesPage/8605/TNO1.pdf>

PRECISION CONTROL OF A QUADRUPLE PENDULUM TEST MASS SUSPENSION FOR THE LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY (LIGO)

Brett N. Shapiro
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

1 INTRODUCTION

Gravitational waves (GWs), predicted to exist by Einstein's General Theory of Relativity, are distortions of space-time that, due to their weak interactions, until now have not been directly detected. Over the next several years, the Laser Interferometer Gravitational-Wave Observatory (LIGO) and other highly sensitive gravity-wave observatories will undergo upgrades that should make detections of GWs possible [1].

We have developed an active control system for the quadruple pendulum, from which the mirrors of the upgraded LIGO interferometers are suspended. This pendulum is a chain of four masses used to provide seismic isolation of the mirrors at the level of $10^{-19} \text{ mHz}^{-1/2}$ at 10 Hz. At this sensitivity level, the sensor noise used in the active control is not trivial. Our goal is to optimize a control scheme that has high gain at the resonant frequencies of the pendulum to provide damping while at the same time rolling the gain off to virtually zero at the limit of the gravitational wave detection band less than half a decade away. This requirement is very difficult to achieve with classical loop shaping design techniques [2].

In addition to this damping control, the position of the bottom mass mirror in the pendulum must be controlled to track the position of another similar mirror 4km away to within an average precision of $10^{-14} \text{ m}_{\text{rms}}$ [3]. At this level of precision, nonlinear noise sources and disturbances are not trivial. Consequently, the control action for the mirror position is split between all four stages of the quadruple pendulum, in what is known in LIGO as the global hierarchical control. The large lowest frequency forces, which contain the highest amplitude noise, are applied to the top mass of the pendulum where they are filtered by the pendulum chain below. Each next stage below is then given a frequency band slightly higher than the stage above it. The bandwidth at each

consecutive stage is limited by the strength of its actuators, the phase loss with respect to the error signal at the bottom mass, and the magnitude of motion at the suspension point of the pendulum.

2 QUADRUPLE PENDULUM ("QUAD") DESIGN

Figure 1 provides an illustration of the quad. Figure 2 shows the coordinate system which will be referenced throughout this paper. The quad consists of two vertical chains of four masses. The masses on each chain are numbered 1 through 4, where the fourth mass on the main (front) chain functions as the optical component. In each chain the top two masses are approximately 20 kg and the lower two masses 40 kg. The chains are about 2 m from top to bottom and hang 5 mm from each other. Each stage of the pendulum provides f^2 isolation above the pendulum's resonant frequencies. Thus, by using four masses a performance of f^8 is achieved. The reaction (back) chain is used to provide a quiet actuation platform to filter any disturbance or noise that might couple through the actuators.

There are six collocated sensor actuator devices called OSEMs (Optical Sensor Electro-Magnet) placed around each top mass and referenced to the ground. These are used to damp the mechanical resonances of the quad. The mechanical resonances, all between 0.5 Hz and 5 Hz, are purposely designed with very high quality factors to optimize the spectrum of thermally excited stochastic motions of the mirror. Damping control of the quad is permitted only at the top mass since the OSEM sensor noise is non-negligible compared to the high sensitivity required by LIGO. As a result, the quad mechanically attenuates the sensor noise through the pendulum chain below. All modes of vibration are designed to couple to the top mass to ensure controllability for the damping loops [4].

The position and angular control of the mirror is done with actuators placed between the main and reaction chains at the second, third, and fourth stages down. At the second and third stages there are four OSEMs and at the bottom stage there are four electrostatic actuators known as the Electrostatic Drive (ESD). The error signal sent to each of these stages is the position of only the bottom stage measured directly from the LIGO interferometer. The OSEM sensors are not used here. The top mass OSEM actuators may also be used if more actuation is desired.

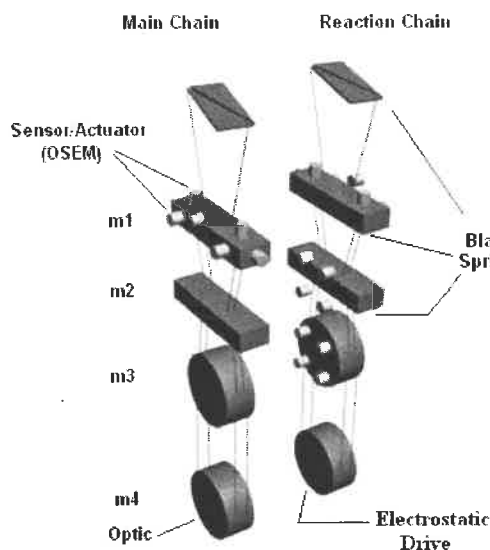


FIGURE 1. Illustration of the quad pendulum. The quad consists of two chains, a main chain and a reaction chain. The bottom mass of the main chain is the interferometer optic. Three stages of blade (cantilever) springs provide vertical isolation. Sensor actuator devices (OSEMs) provide active damping and control. The reaction chain is used as a quiet actuation surface to reduce noise injected by the OSEMs.

FIGURE 2. The orientation of the quad pendulum coordinate system relative to the quad top mass. The positive x-axis points in the direction of the bottom mass mirror.

3 MODAL DAMPING CONTROL

Each quad chain has 24 mechanical resonances between 0.5 Hz and 5 Hz that must be damped without injecting any sensor noise at 10 Hz. Further, this damping control may not couple the quad to the ground as the purpose of the quad is to decouple the mirror from the ground. Consequently, we would like to have damping control with high gain at the resonances, but no gain at 0 Hz and greater than 10 Hz. Achieving zero gain at 0 Hz is easy enough by adding an AC coupled zero to each damping filter. However, the zero gain at 10 Hz is not possible, but we can get sufficiently close with a careful design.

To simplify the control design the 24 coupled degrees of freedom (DOFs) relative to the coordinate system given in Figure 2 the quad is mathematically decoupled into 24 DOFs in a modal coordinate system. Thus, the pendulum is broken down into a series of second order single DOF systems resonant at the pendulum's modes. Each modal plant behaves like a simple mass-spring oscillator where the modal stiffness and modal mass determine the respective resonant frequency. Then it is sufficient to use one simple design of a damping filter and adapt it to each of these individual one DOF systems by a proportional shift of the poles and zeros in magnitude and frequency (horizontal and vertical shift in the Bode magnitude plot). The gains on each filter can be tuned to provide just enough damping so that a sufficiently small amount of noise is introduced. Since the lower frequency modes inject the least sensor noise and, in general, contribute the most mechanical energy, the gains will tend to be set higher on them.

3.1 Modal Decomposition

The equation of motion of the pendulum is represented in Eqn. 1 below.

$$M\ddot{x} + Kx = P \quad (1)$$

M is the mass matrix, K the stiffness matrix, P the vector of control forces and x the position vector in the pendulum coordinate system. The transformation of x into the modal coordinate system is the basis of eigenvectors Φ .

$$x = \Phi q \quad (2)$$

Applying this transformation and multiplying on the left by the inverse of the eigenvector matrix we get the resulting new equation of motion below.

$$M_m \ddot{q} + K_m q = P_m \quad (3)$$

The subscript m references the modal transformation of the mass and stiffness matrices as well as the control forces. Note that now these matrices are diagonal, which is a mathematical result of the fact that the dynamics of each modal DOF are decoupled and can be considered independently.

3.2 Damping Application

For the application to damping control of a real system there are two important transformations to keep in mind.

$$q = \Phi^{-1} x \quad (4)$$

$$P = (\Phi^T)^{-1} P_m \quad (5)$$

Eqn. 4, transforms the sensor signals into the modal signals. Control can then be applied in the modal coordinates where the dynamics are decoupled, which then results in the modal actuation forces P_m . The second transformation, Eqn. 5, converts the modal forces into physical forces that can be applied to the pendulum by the actuators.

A simple graphical representation of this process is provided in Figure 3 below.

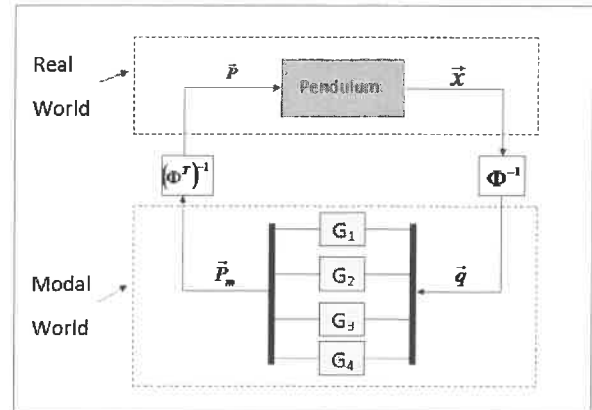


FIGURE 3. A block diagram a modal damping scheme. The coordinate vector is transformed into a modal vector through the inverse of the eigenvector matrix. The modal signals are then sent to damping filters, one for each DOF. The resulting modal damping forces are brought back into the pendulum coordinate system through the inverse of the transpose of the eigenvector matrix. Note that this figure applies to a four DOF system since only four damping filters are shown.

The nominal design of the damping filters is relatively simple since each modal DOF is essentially identical except shifted in frequency and magnitude. Figure 4 shows the open loop gain transfer function for a 1 Hz modal oscillator.

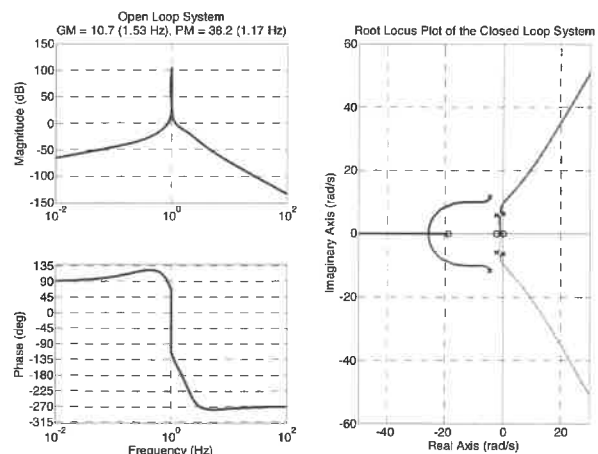


FIGURE 4. Left, the loop gain transfer function of a 1 Hz modal oscillator with its damping filter. Right, the root locus plot of the modal damping loop. The plant contributes the pole pair on the imaginary axis, the damping filter contributes the remaining poles and zeros. All the damping loops have the same shape but shifted in frequency and gain.

3.3 State Estimation

Note that the mathematics of modal damping require the positions of all four masses to be measured. However, only the top mass is directly observed. The stages below have effectively no sensors since any measurement would reference a moving platform with its own dynamics. Consequently, an observer must be employed to reconstruct the full dynamics.

Similarly, actuation is only available at the top mass for this damping control. Although actuators are available at the lower stages, employing them would allow part of the sensor noise to bypass the mechanical filtering of the stages above. Fortunately, a projection of the modal damping forces to the subset of top mass actuation still results in stable control, merely with some reduced damping ability.

The observer separation principle allows us to keep the already established control scheme provided the observer has an accurate model of the pendulum, which is the case here. The main change in application is that the observer will replace the signal transformation from the pendulum coordinates to the modal coordinates.

The estimator is designed using the LQR technique where the cost function is defined in Eqn. 6. The Q matrix weights the accuracy of the

$$J = \int_0^{\infty} (\tilde{q}^T Q \tilde{q} + z_m^T R z_m) dt \quad (6)$$

modal state estimation while the R matrix weights the cost of using a noisy measurement. z_m is defined as

$$z_m = -L_m^T \tilde{q} \quad (7)$$

Where L_m is the estimator feedback matrix determined by the cost function in (6).

A successful observer could be generated in the form of a Kalman filter for this application. However, a Kalman filter is concerned with optimal sensor data recovery whereas our goal is to optimize the noise reaching the bottom mass of the pendulum. Sacrificing sensor signal accuracy is acceptable to a certain degree if the noise performance is improved. Consequently, a more direct approach is taken to obtain the values of Q and R .

First, Q is set by placing the square of the inverse of the resonance frequencies on the diagonal. In this way lower resonance frequencies, which have more mechanical energy, will be damped more efficiently. The value of R only remains to be optimized now.

If we restrict R to be diagonal, a common assumption, we are still left with six parameters to simultaneously optimize relating to six measurement signals. To simplify this problem one can invoke the symmetry of the pendulum. For example, all the modes that represent vertical bounce motion along the z -DOF are decoupled from the others. Thus, these modes can be thought of as representing a separate system which has only four DOFs, and a single sensor signal. The same decomposition can be made along the x -DOFs and yaw-DOFs. The y -, pitch-, and roll-DOFs turn out to be inseparably coupled.

Here we illustrate the optimization process with one of the simpler systems, yaw. Since both the control and the Q matrix are set, the performance of the system is simulated for many values of the scalar R over a sufficiently large interval. The expected spectrum of the suspension point (ground) motion and sensor noise is included in the simulation. The two performance criteria are the settling time of the test mass motion (Figure 5), and the amplitude spectral density of the motion at 10 Hz and above. Since the amplitude spectral density rolls off steeply above 10 Hz, we will consider only the highest value at 10 Hz (Figure 6).

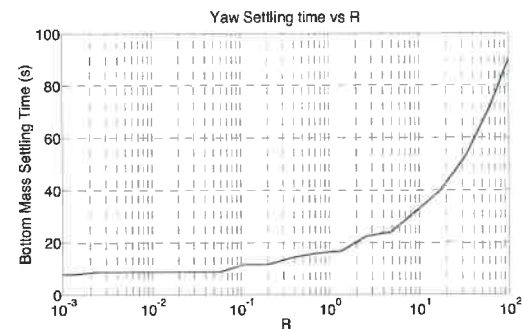


FIGURE 5. The settling time of the bottom mass in yaw versus the cost function parameter R . Small values of R approach the limit where the full state is completely observed. Large values of R approach the limit where there is no measurement.

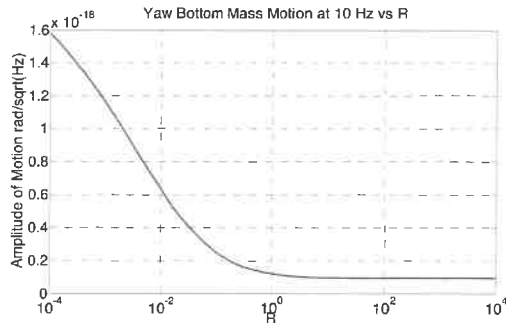


FIGURE 6. Amplitude spectral density of the bottom mass yaw motion at 10 Hz as a function of R . As R gets larger the limit of zero sensor noise injection is achieved.

An R value of 0.1 is chosen to optimize both curves.

The optimization process is similar for the other separated pendulum systems. However, the more complex y-pitch-roll system is not so easily plotted since the optimization has three independent variables. In this case a space of R is sampled within a reasonable interval and the computer is asked to simply choose the minimum of a new cost function, Eqn. 8.

$$J_R = T_s^2 + N^2 \quad (8)$$

T_s is the normalized bottom mass settling time; chosen as the maximum of y , pitch or roll; and normalized by the minimum settling time. N is the normalized amplitude spectral density of the bottom mass x-DOF at 10 Hz as before, but normalized by its minimum value. The computer is essentially doing automatically exactly what was done by eye in Figures 5 and 6.

For a basis of comparison, the yaw performance is analyzed for damping loops designed with this modal damping technique and with typical loop shaping methods in Figure 6. The loop shaped design is somewhat arbitrary and could in theory be optimized given sufficient time. However, we are interested in designing optimized loops without spending arbitrarily long times on the design process, especially for the more complicated coupled DOFs. The loop shaped design in Figure 7 was done by a knowledgeable controls engineer. For similar damping performance, the modal damping loop provides about two orders of magnitude more rejection above 10 Hz.

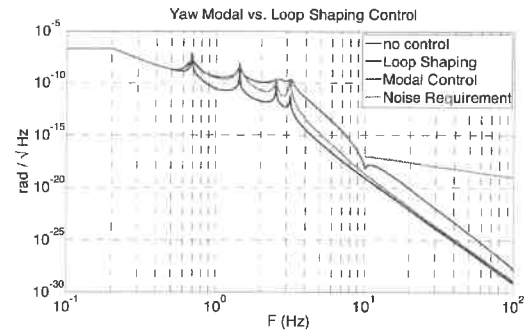


FIGURE 7. An amplitude spectrum comparison of loop shaping to modal damping noise rejection for the yaw-DOF. The blue curve is the performance with the loop shaping designed loop, the red is with modal damping designed loop, and the black is with no damping at all (the theoretical noise limit). The magenta curve is the requirement. Modal damping provides about two orders of magnitude more rejection above 10 Hz for similar damping performance.

4 MIRROR TRACKING CONTROL

In order for the LIGO interferometer to stay locked at its operating point the differential motion between the 4km spaced mirrors must be controlled to within 10^{-14} m_{rms}. Feedback control is employed to maintain this requirement in the presence of seismic disturbances and other forms of process noise that have propagated through the pendulum.

The dynamic range of the actuators was designed to decrease as one moves down the pendulum chain to limit process noise injected into the mirror motion through the control loops. Consequently, one must distribute the control effort of this locking control between the stages of the pendulum. Generally smaller, higher frequency forces are applied at lower stages close to the mirror while larger, lower frequency forces are applied at the higher stages.

The reason relatively small forces are needed at higher frequencies is a result of two facts. First, the seismic disturbance rolls off rather quickly with frequency. Second, and more importantly, actuation closer to the DOF being controlled results in less phase loss and consequently a shallower roll off in magnitude above the resonance frequencies. Thus, there is more mechanical advantage pushing on the pendulum closer to the point of interest, as one might expect.

Since the sensor signal is the GW signal itself, sensor noise is not enhanced by this controller. Despite the limited dynamic range of the actuators, the controller is assumed to contribute process noise, providing further motivation for moving the control effort up the pendulum chain as much as possible.

This type of control distribution is known in LIGO as hierarchical control. Simulation results are plotted in Figures 8 and 9 to illustrate the method.

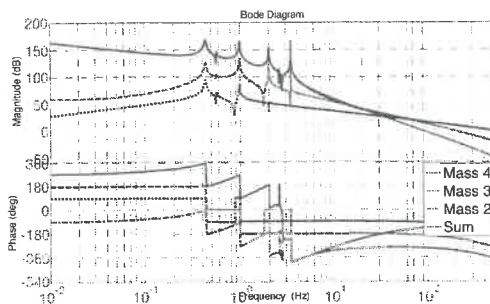


FIGURE 8. The loop gain transfer function for an example hierarchical control scheme to maintain lock at the operation point of the LIGO interferometers. The loop gain component to each stage actuated upon is included along with the overall summed loop gain.

Figure 8 shows the loop gain transfer function along with its components for each stage of the pendulum in the hierarchical scheme. Since this is a multi-input-single-output (MISO) controller the overall loop gain is found by summing the loop gain for each input. The overall loop gain follows the same Nyquist stability criteria as any other SISO loop gain transfer function. However, there are similar gain and phase stability constraints between the component loop transfer functions as well. Note for example that the loops are within 180 degrees of phase when they cross each other while carrying the sum.

Figure 9 shows a simulation of the error signal given the expected seismic disturbance. The RMS limit of the error signal achieved by this simple controller is less than half the requirement. The actuators use at most 10% of their full range. Further optimization is possible since the top mass is not even being used in this loop.

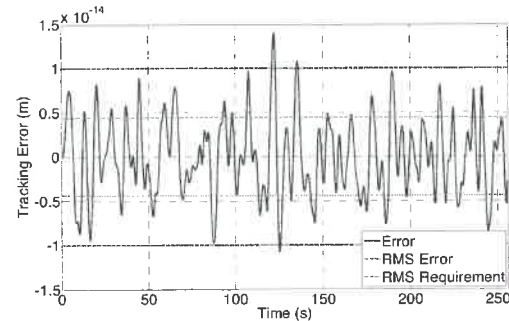


FIGURE 9. A plot of the error signal and its RMS value relative to the requirement. The RMS value achieved here is less than half the requirement.

5 CONCLUSION

To achieve the high sensitivity performance of the LIGO interferometers quadruple pendulums are employed to isolate the mirrors from seismically induced ground motions. In order to hold these mirrors at the interferometer operating point control loops are used to damp the resonance frequencies and force the mirrors to follow each other within a relative displacement limit. A modal damping scheme is chosen to optimize the damping versus sensor noise trade off. A hierarchical scheme is chosen to solve the tracking problem. More detailed models of the process noise involved with the hierarchical scheme are needed to precisely estimate the overall noise performance.

REFERENCES

- [1] Hough J, Rowan S, Sathyaprakash B. The Search for Gravitational Waves. *J. Phys. B: At. Mol. Opt. Phys.* 2005; 38: S497–S519
- [2] Ruet L. Active Control and Sensor Noise Filtering Duality Application to Advanced LIGO Suspensions. Institut National des Sciences Appliquées de Lyon. 2007
- [3] Shoemaker D. Advanced LIGO Reference Design. LIGO Internal Document. 2009; M060056:
<https://dcc.ligo.org/DocDB/0001/M060056/01/AdL-reference-design-v1.pdf>
- [4] Robertson N, et al. Quadruple Suspension Design for Advanced LIGO. *Classical and Quantum Gravity.* 2002; 19: 4043-4058

PROTOTYPING, TESTING AND PERFORMANCE OF THE TWO-STAGE SEISMIC ISOLATION SYSTEM FOR ADVANCED LIGO GRAVITATIONAL WAVE DETECTORS

Fabrice Matichard^{1,2}, Richard Mittleman¹, Brian Lantz³, Ben Abbott², Sam Barnum¹, Dennis Coyne², Dan DeBra³, Stephany Foley¹, Joseph Giaime⁴, Corey Gray⁵, Joe Hanson⁴, Jeff Kissel⁶, Myron MacInnis¹, Ken Mason¹, Brian O'Reilly⁴, Hugh Radkins⁵, Céline Ramet⁴, David Shoemaker¹, Andy Stein¹.

¹LIGO, MIT
Cambridge, MA, US.

²LIGO, Caltech
Pasadena, CA, US.

³Stanford University, Ginzton
Laboratory, Stanford, CA, US.

⁴LIGO Livingston Observatory
Livingston, LA, US.

⁵LIGO Hanford Observatory
Richland, WA, US.

⁶Louisiana State University
Baton Rouge, LA, US.

1. INTRODUCTION

The goal of LIGO is to detect and analyze the gravitational waves produced by astrophysical events [1]. LIGO consists of 4 km scale interferometric detectors located in Washington (LIGO Hanford Observatory, LHO) and Louisiana (LIGO Livingston Observatory, LLO) [2]. Gravitational waves appear as differential length changes in the perpendicular arms. To reach the required level of sensitivity (to motions of 10^{-18} m or less), the interferometer test masses must be isolated from several noise sources, including seismic motion, which is dominant at low frequencies. Initial LIGO has searched for signals, but none have yet been identified.

A significant improvement in sensitivity is underway with the pending installation of Advanced LIGO [3] (or aLIGO). Once running at full sensitivity, sources of gravitational are predicted to be regularly detected. A key element in this upgrade is the installation of extremely effective broadband seismic isolation and positioning systems [4-5]. It will include, for every optic, three levels of isolation in series: an external (in air) active isolation and alignment stage with quiet-hydraulic actuators called HEPI [6], an internal (in vacuum) passive/active isolator called ISI (Internal Seismic Isolator) presented in this paper, and predominantly passive multiple pendulums to hold the optics [7].

A multi-stage Internal Seismic Isolation platform will be used to support the beam splitter and each of the test mass optics. The platform should provide an isolation factor of 10 at 0.1 Hz

and up to 2000 at 10 Hz. This active/passive system, called BSC-ISI, under development for several years, is made of two stages in series, each controlled in six degrees of freedom (DOF).

This paper presents the prototyping, testing and current performance of the BSC-ISI system supporting a 920 kg total payload (370 kg of optical payload with the remaining 550 kg configured as ballast). In the next section, the BSC-ISI architecture and the prototype are presented. The active control strategy is presented in section 3. Initial performance results and related structural investigation are presented in section 4, followed by improved performance in section 5.

2. BSC-ISI ARCHITECTURE AND PROTOTYPE

A rendering of the BSC Internal Seismic Isolation platform can be seen in figure 1. Either the triple pendulum suspension for the beam splitter or the quadruple pendulum suspension for the test masses (neither are shown) would be suspended from the optical table on the bottom of the second stage.

The base of the system is shown in blue on the CAD model in figure 2. It supports Stage 1, shown in cyan, via blades and flexure rods shown in yellow. The blades provide the vertical flexibility and the rods provide the horizontal one. Stage 1 supports Stage 2, shown in grey, via the same type of blades and flexure rods also shown in yellow. The two suspended stages have natural frequencies in the 1 Hz - 7 Hz range. The system is designed to provide passive isolation above its natural frequencies

and active control isolation in the 0.1 Hz - 20 Hz range.

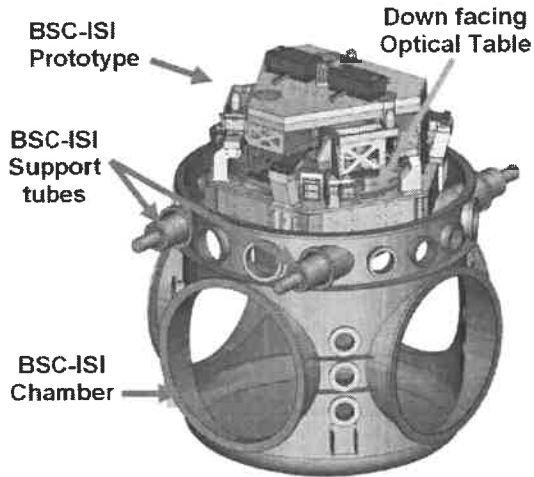


FIGURE 1. Model of the BSC-ISI on its supports tubes at the top of the BSC chamber (doors, dome and flanges not shown)

For the active control, six coarse (high force) electromagnetic actuators are used between the base and stage 1. They are shown in pink on figure 2. Six smaller fine (low force) electromagnetic actuators, also shown in pink on figure 2, are used to actuate between stage 1 and stage 2. The blades, rods and actuators have been designed and positioned to decouple the horizontal and vertical motion of the stages.

Sensors are shown in red on figure 2. Stage 1 is instrumented with three different sets of sensors:

- 6 MicroSense (formerly ADE technologies) capacitive position sensors between the base and stage 1, called "coarse". They are used for low frequency relative positioning.

- 3 Streckeisen STS-2 seismometers: these 3-axis seismometers provide the low frequency inertial measurements necessary to provide active seismic isolation.

- 6 Sercel L-4C geophones: these single-axis seismometers provide the high frequency inertial sensing necessary to provide active seismic isolation at higher frequencies.

Stage 2 is instrumented with two different sets of sensors:

- 6 MicroSense capacitive position sensors between stage 1 and stage 2, called "fine". They provide the low frequency relative positioning.
- 6 Geotech GS-13 seismometers: these single-axis seismometers provide the inertial sensing needed for active seismic isolation.

The prototype installed on the support tubes at the LIGO MIT test facilities is shown in figure 3.

The bottom plate of Stage 2 is the down-facing optical table supporting the optics.

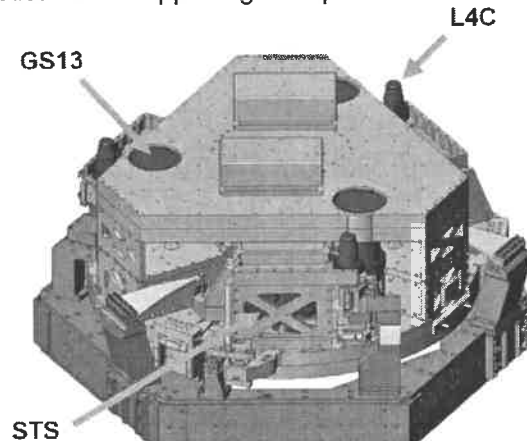


FIGURE 2. CAD model of the BSC-ISI.

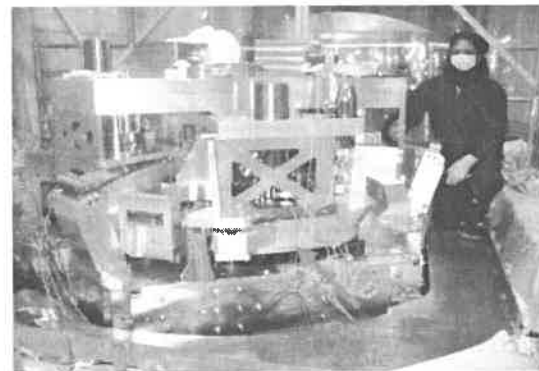


FIGURE 3. The BSC-ISI prototype at LIGO MIT.

3. ACTIVE CONTROL STRATEGY

The feedback control approach is described on the diagram in figure 4.

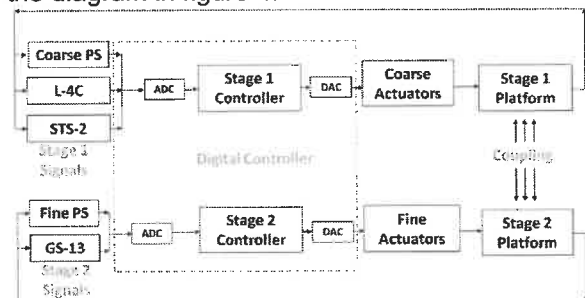


FIGURE 4. Active control general diagram.

It shows the input control signals of stage 1 made of the 6 Position Sensors (PS), 3 three-axis seismometers (STS-2) and 6 single axis seismometers (L-4C). It shows the input control signals of Stage 2 consisting of 6 position sensors (PS), and 6 single-axis seismometers (GS-13).

Those signals are digitized and sent to stage 1 and stage 2 controllers that drive the coarse and fine actuators, respectively. Although stage

motions are coupled to each other, the stages can be controlled independently as explained in the next paragraph which describes the controllers structure.

For each stage, the control is based on the use of 6 independent SISO loops. The control is done in the basis of the General Coordinate System (GCS) of the MIT test facilities: X, Y, RZ, Z, RX, RY. X and Y are aligned with the arms of the interferometer, Z is the vertical axis and RX, RY and RZ the rotations around those axes. The origin of the GCS is on the plan of the coarse horizontal actuators.

Figure 5 shows the block diagram used for each of the 6 digital controllers of each stage. Each SISO controller senses and controls in one direction or one rotation. Firstly, the 6 position sensors of one stage are combined to obtain the motion of this stage in the GCS through a Local to General coordinate change of basis matrix (L2G on block 1 in fig. 5). The same thing is done for the seismometers to get the inertial motion of the stage in the GCS.

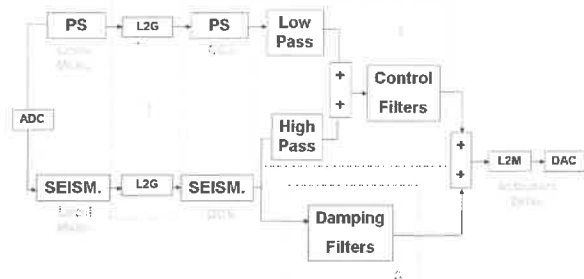


FIGURE 5. Block diagram of one DOF.

The seismometer signal in the GCS is first used to damp the suspension resonances (block 3 on Fig 5). This makes the further developments easier.

Position sensors and seismometers signals are then blended [8]: the position sensors signal is filtered by a low pass and the seismometer is filtered by a high pass. The two signals are summed resulting in a "super sensor" (block 2 on figure 5). The super sensor signal is dominated by relative displacement sensing at low frequency and inertial sensing at high frequency. The frequency at which relative participation and inertial sensing have equal participation is referred to as the blend frequency (BF).

Finally, a control filter is applied to the super-sensor to provide loop gain. These filters typically are designed for a unity gain frequency (UGF) between 20 Hz and 30 Hz. The control strategy locks the stages to the base below the BF, which provides positioning and alignment

capabilities, and drift rejection. The control provides seismic isolation from the BF (around a few tenths of a hertz or below) to the UGF.

4. INITIAL PERFORMANCE RESULTS AND ANALYSIS

Initial performance obtained after the first BSC-ISI commissioning are presented in figure 6. The black curve is the input disturbance (motion of the HEPI system supporting the BSC-ISI). The blue curve shows the motion of stage 2 when the control is off. The system provides passive isolation as expected above 7 Hz. The magenta curve shows the motion of stage 2 with the control on. In this example, the active control improves the isolation between 0.25 Hz and 10 Hz.

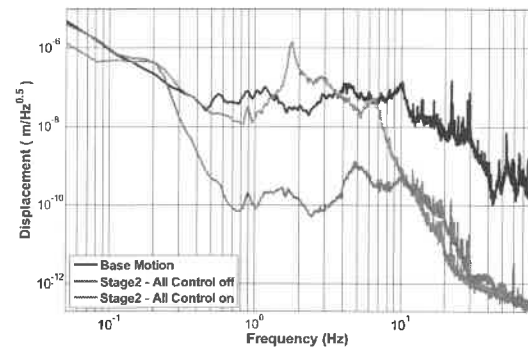


FIGURE 6. Initial performance - Z direction

At low frequencies, one the main factors limiting the performance is the sensitivity of horizontal seismometers to tilt motion. At high frequencies it is limited by the high modal density of the structure and the payload, as seen in the transfer function in figure 7.

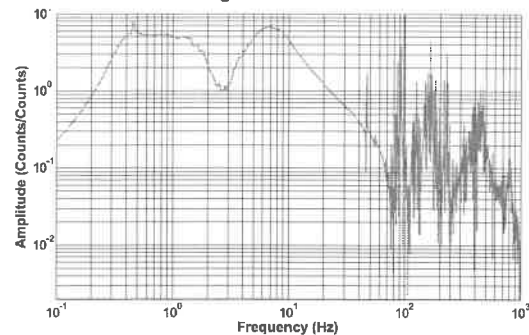


FIGURE 7. Example of transfer functions. Drive X to Sense X with GS13 - Oct 2008.

Initial control results showed that the modal density due to structural resonances needed to be lowered in order to make the control more robust and to improve the performance.

Modal testing investigation identified modes related to the optical payload. The current optical payload under test is a 370 kg system,

whose structural resonances couple to the BSC-ISI. We are currently investigating methods to damp these modes passively. Other sources were identified, such as the counterweights used to balance the system and the seismometer vacuum chambers being mounted improperly resulting in cantilevered masses.

The plot below illustrates an example of structural improvements. The red curve shows the initial transfer functions. The black one shows the transfer functions after the counter weights attachment were re-engineered. These structural modifications significantly improved the control performance as described in section 5.

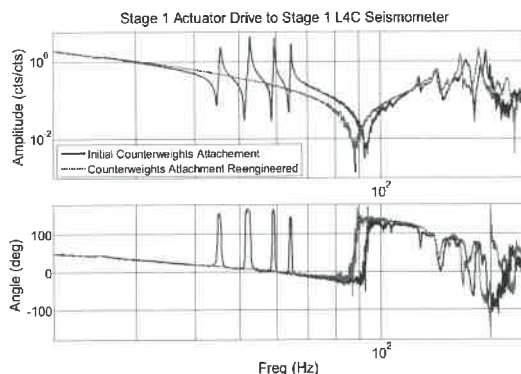


FIGURE 8. Transfer function before and after counterweight attachment reengineering.

5. IMPROVED PERFORMANCE

This section shows the control performance after the structural investigation to improve control robustness and performance. The plot in figure 9 shows the motion of the optical table measured in the X direction with the GS-13s. The black curve shows the motion of the table when the control is off. The blue curve shows the motion of the table under High Frequency Blend (HFB) control on both stages, which is an intermediate control step with reduced isolation. Position sensors and seismometers are blended at 0.7 Hz and the active control provides isolation from 1 Hz to 20 Hz.

Once the system is under HFB control, the "X to RY" and "Y to RX" low frequency couplings induced by position sensors misalignment can be evaluated and minimized. After decoupling, there is less tilt induced by horizontal motion and the blend frequency can be lowered to put the system under Low Frequency Blend (LFB) control. The red curve shows the results after the BF has been lowered to 0.2 Hz on stage 2. The active control provides isolation from 0.25 Hz to 20 Hz.

The purple curve shows the results of implementing a sensor correction. The principle of sensor correction is to sum stage 1 inertial sensing with stage 2 relative displacement sensing, resulting in a low frequency inertial sensing of stage 2. This reduces the gain peaking at the blend frequency.

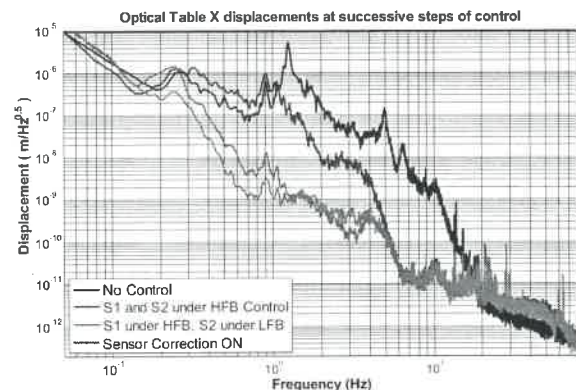


FIGURE 9. Improved performance - X direction.

The next step will be to lower the blend frequency on stage 1 which should provide the same level of improvement as from the blue curve to the red curve. At this stage of commissioning, the system meets isolation expectations for this intermediate level of control:

- At low frequencies (around 0.1 Hz) there is very little motion amplification. This is where there is a risk of gain peaking induced by the relative displacement signal part of the control loop. The sensor correction reduces this gain peaking.
- The isolation starts as low as 0.1 Hz. The tilt in horizontal seismometers will have to be reduced in order to further lower the BF.
- The isolation at 1 Hz is close to a factor of 100. Higher isolation would be at the cost of a higher gain peaking at the BF.
- The unity gain frequency is designed to be 30 Hz. The control provides isolation up to 20 Hz. The phase margin is small in order to maximize the loop gain near 10 Hz (the gain is nearly 10 at 10 Hz). The drawback is higher gain peaking above 20 Hz. This is a good trade-off since the ISI must provide more isolation at 10 Hz than at 20 Hz and above, where the quadruple pendulum provides excellent passive isolation for the final optic.

The same steps have been followed in all of the degrees of freedom and provide similar performance in Y, Z and RZ.

A slightly higher blend frequency is used for the RX and RY rotations around horizontal axes. The RX control performance is presented in figure 10 to illustrate the rotation control performance. The isolation characteristics are similar to those previously described in the other degrees of freedom except for the blend frequency. In this case the BF is set at 0.4 Hz. This provides isolation from 0.5 Hz and above. The reason why the blend frequency is set higher on the rotation control is to reduce the low frequency gain peaking and consequently the tilt motion amplification that would couple with the control of the horizontal seismometers.

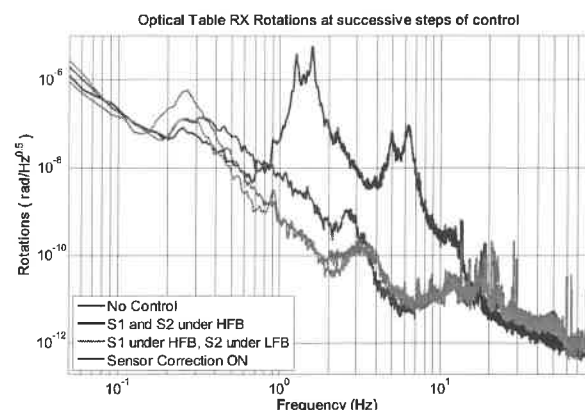


FIGURE 10. RX control performance.

In the previous plots, the control performance was presented (performance with control on versus performance with control off). In the following, we present the overall performance by comparing the motion of the ground with the motion at the top of the piers (BSC-ISI input), the motion of the optical table with control off and the motion of the optical table with control on. The requirement curves are based on aLIGO requirements normalized for the noisier seismic environment of the LIGO MIT urban location.

The performances in the X direction are presented in figure 11:

- The ground motion is shown in black. It is measured with a STS-2 seismometer sitting on the ground 20 feet away from the BSC chamber.
- The motion at the top of the piers is shown in blue. It is measured with HEPI L4Cs.
- The motion of stage 2 when the control is off is presented in red
- The motion of stage 2 when the control is on is presented in purple
- The requirements to be met at MIT in order to meet aLIGO requirement at LLO are presented in grey.

- The requirements to be met at MIT in order to meet aLIGO requirement at LHO are presented in green.

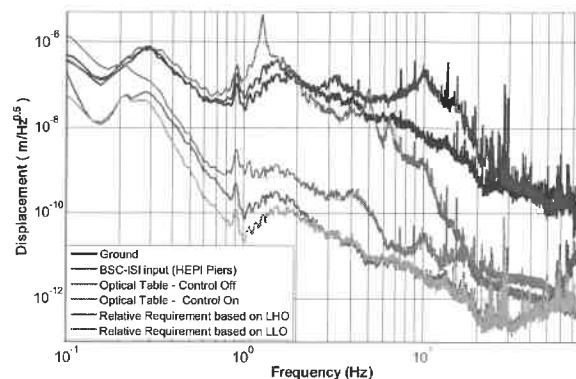


FIGURE 11. BSC-ISI performance – X direction

The blue curve is the input motion of the BSC-ISI system. It is interesting to note that the motion at the top of the chamber is significantly higher than the ground motion. This motion amplification around 10 Hz is due to the motion of the piers driven by the rocking of the vacuum chamber that is heavy enough to rock the ground.

The red curve shows the motion of the optical table when the active control is off. The attenuation from the blue curve to the red one shows the BSC-ISI passive isolation which provides an isolation factor of 100 at 10 Hz and 800 at 20 Hz. The motion of the optical table when the control is on is shown in purple. The system provides a total isolation factor of 6800 at 10 Hz.

The HEPI control and the lower frequency blend that will be implemented on stage 1 is expected to help to meet the requirements from 0.2 Hz to about 8 Hz.

The sensor correction from stage 1 to stage 2 will be adjusted in order to improve the performances around 0.1 Hz. Above 10 Hz it will be difficult to meet the initial requirements. This is not because the BSC-ISI doesn't meet the performance expected by design. It is due to the high input motion at the top of the piers. This is not a critical issue since a recent aLIGO overall noise budget has shown that the seismic noise should be below other sources of noise at those frequencies.

The performance in the vertical direction is presented on figure 12. Like for the horizontal directions, the HEPI control and stage 1 lower frequency blend should help meeting requirements up to about 16 Hz.

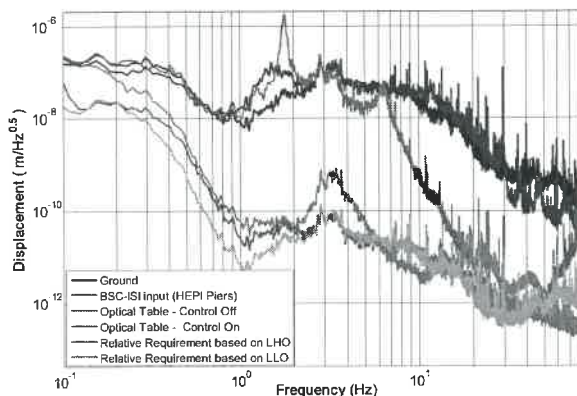


FIGURE 12. BSC-ISI performance – Z direction.

Finally, the plot on figure 13 shows the evolution of the control performance. Each isolation curve consists of the optical table motion when the control is on over the motion when the control is off.

Initial isolation performance, as described in section 4, is shown in black. Current performance is shown in magenta. The slight loss of performance from 0.3 Hz to 0.8 Hz and from 2 Hz to 3 Hz should be largely compensated by the stage 1 blend frequency lowering. This plot shows that the performance around 10 Hz has been improved by a factor of more than 50. It also shows the fine control tuning that is required to minimize the low frequency gain peaking.

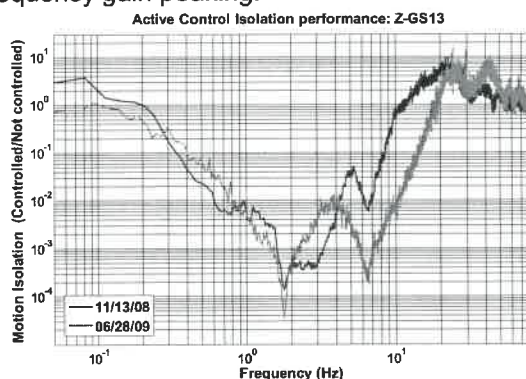


FIGURE 13. Active isolation improvement.

CONCLUSION

The architecture and prototype of the BSC-ISI system, double stage isolation platform that will be used for aLIGO, have been presented. The active control strategy has been described. We presented recent performance measurements of the isolator supporting a 920 kg total payload.

Special attention has been given to the direct correlation between structural modal deformations and the performance achieved. The hardware testing, modal identification and

structural reinforcements made on the prototype were described.

Current performance shows that the BSC-ISI is expected to provide the level of passive and active isolation required by aLIGO. The next steps will be to enhance the performance at low frequencies by limiting the horizontal seismometers tilt coupling and at high frequencies by pursuing the mitigation of structural resonances.

ACKNOWLEDGMENTS

LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0107417. This document has been assigned LIGO Laboratory document number LIGO-P1000029.

REFERENCES

- [1] B Barish, R Weiss. LIGO and the Detection of Gravitational Waves. *Physics Today*. 1999; 52: 44-50.
- [2] B Abbott et al. LIGO: the Laser Interferometer Gravitational-Wave Observatory. *Rep. Prog. Phys.* 2009; 72 No 7: 25pp.
- [3] G. M. Harry et al. Advanced LIGO: The next generation of gravitational wave detectors. *Classical and Quantum Gravity*, to be published in the Proceeding of the Eighth Edoardo Amaldi Conference on Gravitational Waves (2010).
- [4] N A Robertson et al. Seismic isolation and suspension systems for Advanced LIGO. *Gravitational Wave and Particle Astrophysics Detectors*, Proceedings of SPIE. 2004; 5500: 81-91.
- [5] R Abbott, et al. Seismic isolation enhancements for initial and advanced LIGO. *Class. Quantum Grav.* 2004; 21: 915-921.
- [6] C. Hardham et al. Multi-DOF Isolation and Alignment with Quiet Hydraulic Actuators. *Spring topical meeting on Control of Precision Systems*, Proceedings of ASPE. 2004; 32: 127-132.
- [7] N A Robertson, et al. Quadruple suspension design for Advanced LIGO. *Class. Quantum Grav.* 2002; 19: 4043-4058.
- [8] W. Hua et al. Low Frequency Active Vibration Isolation for Advanced LIGO. *Gravitational Wave and Particle Astrophysics Detectors*, Proceedings of SPIE. 2004; 5500: 194-199.

OPTIMIZED DYNAMIC DECOUPLING IN SCANNING STAGES

Marcel F. Heertjes, Arjan J.P. van Engelen, and Maarten Steinbuch
Department of Mechanical Engineering
Eindhoven University of Technology
Eindhoven, The Netherlands

ABSTRACT

To improve performances of multi-input multi-output (MIMO) scanning stage systems, the possibilities of dynamic decoupling are studied. Atop a course static decoupling, a MIMO finite impulse response (FIR) decoupling structure is added for machine-specific and performance-driven fine tunings. The coefficients of the FIR filters are obtained through data-based optimization while the machine operates in closed loop.

INTRODUCTION

In the field of motion control, fast positioning and nano-accurate stages are used for example in wafer scanners [1] and scanning probe microscopes [2]. In the simplified representation of Figure 1 course and micrometer positioning is obtained through long-stroke control whereas fine

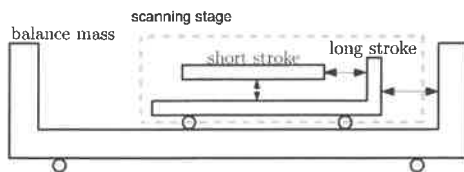


FIGURE 1. Graphical representation of a scanning stage system.

and nanometer positioning is done at the short-stroke scanning stage. The scanning stage system generally interacts with several other subsystems. For example, the displayed balance mass which is used to minimize the effect of the reaction forces generated in the stage system while performing scanning tasks.

Scanning is usually done in the y -direction of an orthogonal basis (x , y , r_z , z , r_x , and r_y) centered about the center-of-gravity of the short-stroke stage mass. To dictate motion in this scanning direction a centralized control design [3] is used to translate the center-of-gravity forces: F_x , F_y , M_{rz} , F_z , M_{rx} , and M_{ry} to actuator forces; see for example Figure 2 for a representative actuator configuration. The underlying actuator and sensor transformation matrices are part of the static decoupling design enabling single-input single-

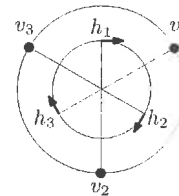


FIGURE 2. Stage actuator configuration.

output (SISO) control of the generally multi-input multi-output (MIMO) scanning stage plant.

The validity of such a SISO control approach is shown in Figure 3. In Bode magnitude represen-

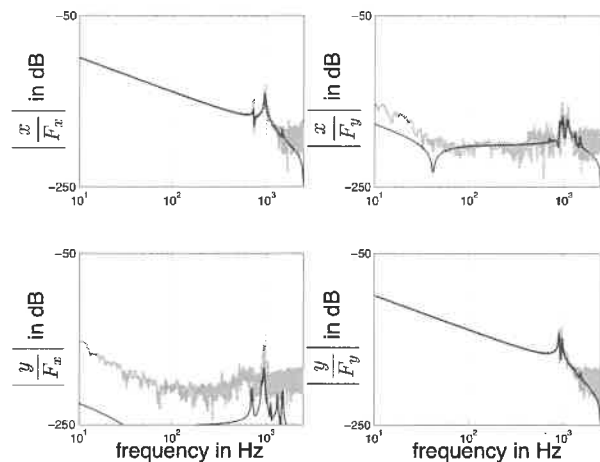


FIGURE 3. Bode magnitude representation of a stage plant; 2×2 intersection of a 6×6 plant.

tation this figure depicts the frequency-domain characteristics of a scanning stage either measured (gray) and modelled (black). Though diagonally dominant clearly the scanning stage has MIMO features: a force in y -direction induces a response in x -direction. This impacts both stability and performances of the control design [4].

Strictly speaking, tuning and loop-shaping of the individual control loops no longer guarantees robust stability and performances and requires a more complex MIMO control design [5]. Moreover, performance is limited because of cross-

coupling: actuator commands and noises injected in one control axis are distributed among the system through the remaining control axes. This is

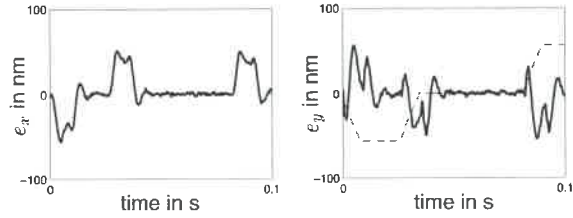


FIGURE 4. The effect of cross-coupling in time-series measurement for an y -scanning set-point.

clearly shown in Figure 4 through measured errors e_x and e_y resulting from a third-order set-point in y -direction. To reduce the effect of cross-coupling and to improve the conditions underlying the SISO control approach the possibilities of MIMO dynamic decoupling are studied [6].

The remainder of the paper is organized as follows. First, a dynamic decoupling using finite impulse response (FIR) filters is considered. Second, the FIR filter coefficients are obtained from a data-based and extremum seeking optimization approach [7]. Third, a frequency-domain measure to quantify the degree of (de)coupling is discussed. Fourth, time-domain performance is assessed on a scanning stage, and, fifth, the pros and cons of optimized dynamic decoupling in the context of scanning stages are reviewed.

DYNAMIC DECOUPLING CONTROL DESIGN

The dynamic decoupling control design stems from the simplified control scheme in Figure 5. $C_{fb}(z)$ represents a 6×6 SISO feedback controller

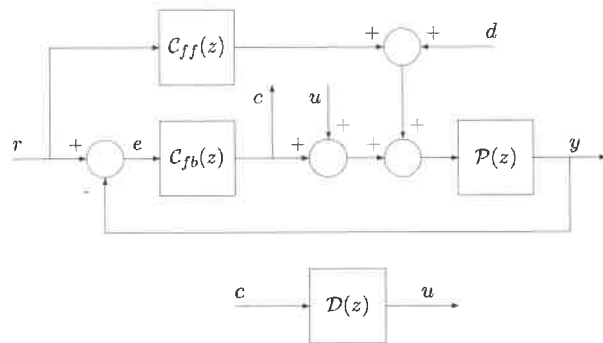


FIGURE 5. Block diagram of the centralized control design of a scanning stage system.

in z -domain, $C_{ff}(z)$ a 6×6 SISO feed-forward controller, and $P(z)$ the 6×6 MIMO scanning stage plant, see [8]. Input to C_{fb} is the (time-domain)

error signal $e(t)$ obtained by subtracting the plant output signal $y(t)$ from the scanning reference commands $r(t)$. Disturbance forces acting on the stage dynamics are represented by $d(t)$.

Dynamic decoupling is now obtained from the extra feedback part in Figure 5 and given by the decoupling transfer $D(z)$ between input (controller output) $c(z)$ and output (plant input) $u(z)$. A proper choice for $D(z)$ follows from the next argument. Consider the desired plant

$$\mathcal{P}^d(z) = \begin{bmatrix} \mathcal{P}_x(z) & & 0 \\ & \ddots & \\ 0 & & \mathcal{P}_{rz}(z) \end{bmatrix}, \quad (1)$$

which only contains the diagonal entries of $\mathcal{P}(z)$. By choosing

$$D(z) = \mathcal{P}^{-1}(z)\mathcal{P}^d(z) - \mathbf{I}, \quad (2)$$

it follows from Figure 5 that the transfer from $c(z)$ to $y(z)$ with $r(t) = d(t) = 0$ reads

$$\frac{y(z)}{c(z)} = \mathcal{P}(z)(\mathbf{I} + D(z)) = \mathcal{P}^d(z), \quad (3)$$

hence a decoupled plant. So in constructing the decoupling $D(z)$ and the desired decoupled plant dynamics $\mathcal{P}^d(z)$ we essentially need the inverted plant dynamics $\mathcal{P}^{-1}(z)$, see also [9].

For both the modelled and measured plant in Figure 3, a 2×2 intersection of the thus obtained dynamic decoupling D is shown in Figure 6. In

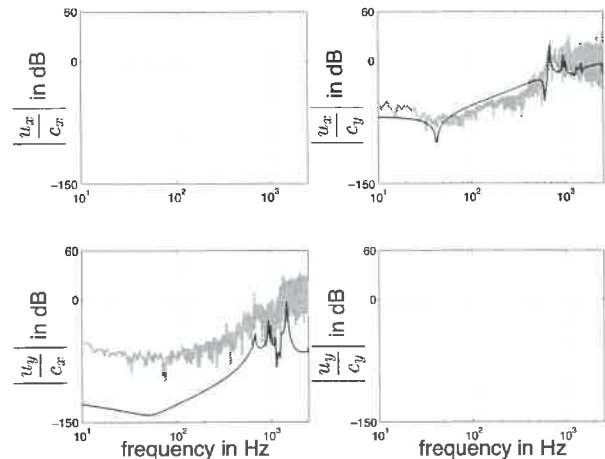


FIGURE 6. Bode magnitude representation of perfect decoupling; 2×2 intersection of 6×6 D .

Bode magnitude representation (and despite the poor quality of measurement) it shows three important features. First, D contains off-diagonal

entries only. Second, up to 600 Hz the frequency response functions are characterized by anti-resonant behavior described by zeros; decoupling based on static gains will not suffice. Third, beyond 600 Hz coupling occurs through resonant behavior described by poles.

In view of these properties $\mathcal{D}$ is designed as

$$\mathcal{D}(z) = \mathcal{F}(z)\mathcal{L}_p(z) \quad (4)$$

which incorporates the off-diagonal structure [10]

$$\mathcal{F}(z) = \begin{bmatrix} 0 & \mathcal{F}_i^j(z) \\ \mathcal{F}_j^i(z) & 0 \end{bmatrix} \quad i, j \in x \quad r_z \quad i \neq j \quad (5)$$

consisting of finite impulse response (FIR) filters. The FIR filter description reads

$$\mathcal{F}_i^j(z) = \theta_0^{ij} + \theta_1^{ij}z^{-1} + \dots + \theta_o^{ij}z^{-o} \quad (6)$$

with o the filter order and θ its coefficients; (6) shows the ability to create zeros using low-order models. The set of low-pass filters is given by

$$\mathcal{L}_p(z) = \begin{bmatrix} \mathcal{L}_{p_x}(z) & & 0 \\ & \ddots & \\ 0 & & \mathcal{L}_{p_{r_z}}(z) \end{bmatrix} \quad (7)$$

and is used to avoid amplification beyond 600 Hz. The reason for including low-pass filters is twofold. On the one hand, low-order FIR filters induce poor high-frequency approximations because of the inability to create poles. On the other hand, high-frequency gains through $\mathcal{D}$ induce high-frequency noise amplification in the control design, which is undesired.

Given the decoupling structure in (4) the question arises how to obtain the FIR filter coefficients θ in (5) and (6). For this purpose we use a data-based optimization approach centered about perturbed-parameter experiments. The motivation is clear: each scanning stage system shows differences in its dynamic behavior that on a nanometer performance scale are too important to neglect. Any model-based approach that does not address these differences simply renders itself ineffective.

DATA-BASED OPTIMIZATION

In finding the FIR coefficients that minimize the effect of cross-talk in scanning stage systems, an extremum seeking optimization scheme is applied, see also [7, 11]. The gradient signals needed in this scheme are obtained through perturbed-parameter experiment [12, 13].

The optimization starts with collecting the relevant cross-talk error signals via a quadratic objective function:

$$V(k) = \mathbf{e}^T(k)\underline{\gamma}\mathbf{e}(k) \quad (8)$$

The servo error signals (denoted with $\mathbf{e}(k)$) result from time sampling in a user-defined interval. This is the interval where we want to achieve performances, hence the scanning interval. $\underline{\gamma}$ is a diagonal weighting matrix having entries $0 < \gamma_1 \dots \gamma_n \leq 1$. It is used to put different weights on the considered cross-talk error signals.

The control objective is to minimize overall cross-talk by minimizing the objective function $V(k)$. We therefore seek the set of FIR filter coefficients in (5) that achieves this goal. A natural way to do so is via the iterative Gauss-Newton scheme [14]:

$$\theta(k+1) = \theta(k) - \zeta (\nabla \mathbf{e}^T(k)\underline{\gamma}\nabla \mathbf{e}(k))^{-1} \nabla \mathbf{e}^T(k)\underline{\gamma}\mathbf{e}(k) \quad (9)$$

with damping constant $0 < \zeta \leq 1$ and k the number of iterations, see also [15].

To obtain the new set of FIR filter coefficients $\theta(k+1)$ two signals are required: the cross-talk error signals $\mathbf{e}$ and the gradient error signals $\nabla \mathbf{e}$. The latter are obtained from perturbed-parameter experiment [16] in which

$$\nabla \mathbf{e}_i = \left. \frac{\partial \mathbf{e}}{\partial \theta_i} \right|_{\Delta \theta_i} \approx \frac{\mathbf{e}_{\theta_i + \Delta \theta_i} - \mathbf{e}_{\theta_i}}{\Delta \theta_i} \quad (10)$$

Herein the gradient error signal with respect to FIR filter coefficient θ_i with $i \in -1 \dots n$ is approximated by doing two closed-loop experiments. In the first experiment all coefficients of all FIR filters are zeroed ($\theta_1 = \dots = \theta_n = 0$). In the second experiment only θ_i is given the value $\Delta \theta_i \neq 0$. Both (9) and (10) render the method strictly data-based: only measured signals are used, neither model nor model knowledge is required. One can argue whether model knowledge is incorporated via the FIR filters/structure in (5).

In Bode magnitude representation Figure 7 shows the effect of perfect decoupling in terms of the considered 2×2 intersection. Additionally the optimized decoupling is shown using the scheme in (9) with $k = 1$, $\zeta = 1$, and $o = 2$. From the figure we conclude that a fairly good approximation is obtained of the low-frequency behavior by minimizing cross-talk error signals under closed-loop conditions. High-frequency resonances, however, are neglected. Having a method to obtain an optimized set of filters $\mathcal{D}(z)$ the question arises how to quantify the degree of (de)coupling.

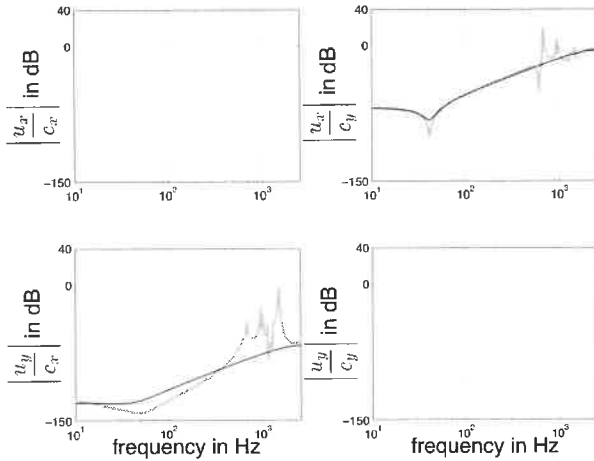


FIGURE 7. Bode magnitude representation of perfect (gray) and optimized (black) decoupling; 2×2 intersection of 6×6 $\mathcal{D}$.

THE DEGREE OF (DE)COUPLING

To quantify the degree of (de)coupling the transfer function matrix $\mathcal{P}^*(z)$ is used as a measure for interaction [17]:

$$\mathcal{P}^*(z) = (\mathcal{P}_o(z) - \mathcal{P}_o^d(z))(\mathcal{P}_o^d(z))^{-1} \quad (11)$$

with the optimized plant $\mathcal{P}_o(z) = \mathcal{P}(z)(\mathbf{I} + \mathcal{D}(z))$ and $\mathcal{P}_o^d(z)$ a diagonal matrix containing only the diagonal elements of $\mathcal{P}_o(z)$. By subtracting $\mathcal{P}_o^d(z)$ from $\mathcal{P}_o(z)$ in (11) only off-diagonal (or interaction) terms are considered for further evaluation.

For the optimized 6×6 plant $\mathcal{P}_o(z)$ the degree of coupling now follows from structured singular value analysis: μ on $\mathcal{P}^*(z)$ using the MatLab function `mu.m`; see also [17]. At each frequency point ω_i the structured singular values $\mu(\mathcal{P}^*(j\omega_i))$ are computed. The largest value is then designated to this frequency point thus giving a single-valued measure on the 6×6 interaction terms.

For the considered simulation model the result of this analysis is depicted in Figure 8. The interpretation of the figure is as follows: diagonal dominance (hence decoupling) occurs for $\mu(\mathcal{P}^*(j\omega)) < 1$. Before optimization and beyond 600 Hz the plant is clearly not diagonally dominant. Optimization neither improves nor deteriorates this. In the low-frequency range, however, the degree of coupling is significantly reduced after optimization; up to 30 dB at 1 Hz!

The same analysis is performed for a scanning wafer stage whose results are shown in Figure 9. The improvement reduces to 8dB. This is be-

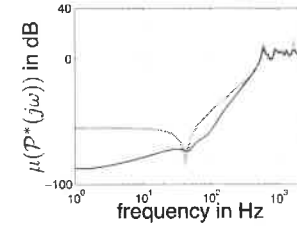


FIGURE 8. Structured singular values in frequency-domain simulation; 6×6 plant $\mathcal{P}^*(j\omega)$ before (gray) and after (black) optimization.

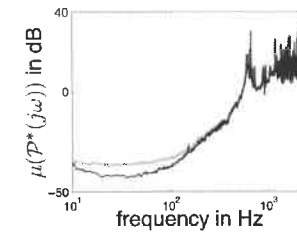


FIGURE 9. Structured singular values in frequency-domain experiment; 6×6 plant $\mathcal{P}^*(j\omega)$ before (gray) and after (black) optimization.

cause of the control aim: minimize the quadratic representation of the cross-talk error signals. The degree of (de)coupling is not directly addresses. Therefore in experiment and subjected to disturbances and system nonlinearities, both aims diverge. In simulation, however, where disturbances are absent and the system model is linear time-invariant both aims seem to match.

STAGE PERFORMANCE ASSESSMENT

To evaluate the effect of the optimization given the objective function in (8) we need to study the cross-talk error signals in time-domain. The effect in y -direction is shown in simulation in Figure 10 (and in measurement in Figure 11). It is clear that

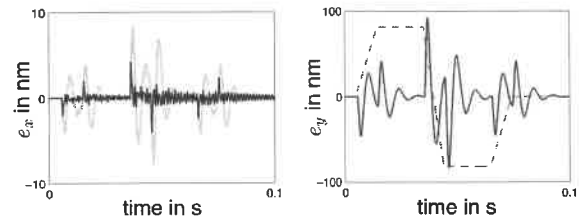


FIGURE 10. Time-series simulation for an y -scanning optimization; 2-dof intersection.

the optimization is effective. Cross-talk error signals prior to optimization (gray) are significantly reduced after optimization (black). In simulation

this improvement comes at no effect at all in the direction of excitation: the y -direction. In measurement even a small improvement can be observed in this direction.

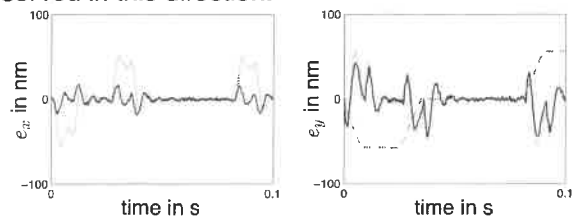


FIGURE 11. Time-series measurement for an y -scanning optimization; 2-dof intersection.

An assessment on stage performance with the optimized decoupling $\mathcal{D}(z)$ is shown in Figure 12. For a standard scanning job, the measured time-

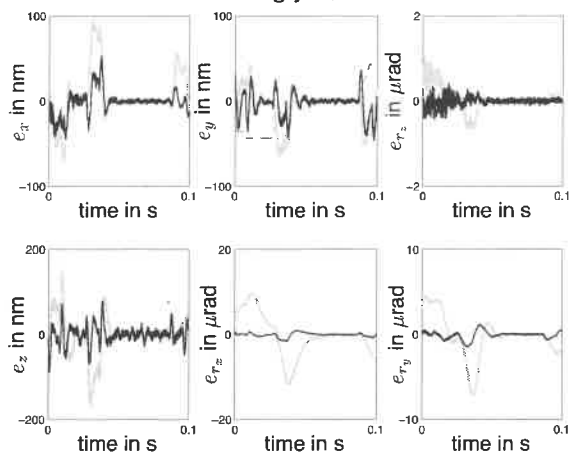


FIGURE 12. Time-series measurement for an y -scanning job; full 6-dof representation.

series are shown before (gray) and after (black) optimization. The improvements (note that every nanometer counts in this industry) involve all six control axes, but reveal one of the key limitations of the approach: the source of cross-talk error signals is not taken into account. As such non-decoupling related effects are potentially addressed by the decoupling. For example, a feed-forward mismatch, closed-loop disturbances, and actuator/amplifier nonlinearities that typically affect the cross-talk error signals are included in the optimization, whether decoupling-related or not.

CONCLUDING REMARKS

In terms of optimized dynamic decoupling the considered data-based approach requires neither model nor model knowledge. This gives rise to a machine-dedicated calibration in which improved scanning stage performances are demonstrated.

Stability of the optimization scheme is not directly imposed, but hints toward a self-stabilizing mechanism: as the closed-loop system tends unstable, the cross-talk errors tend to increase. This increase is avoided by the optimization as soon as it implies an increase of the objective function.

Stability (and similarly convergence [18]) very much depends on the error signals being affine in the FIR filter coefficients. By design, the FIR filter structure has linear dependency on its coefficients. Nevertheless the affine property is only locally valid, i.e., sufficiently close to perfect decoupling. In practice, this means that already a good course decoupling should be present in order for the optimization scheme to converge. For the scanning stage example this is the case: one iteration is required to get near the optimal set of FIR filter coefficients.

In terms of decoupling structure, FIR filters appear advantageous because of their ability to model zeros. Consequently the order of the models can be kept small, because perfect dynamic decoupling of the scanning stage system up to 600 Hz merely requires zeros in its description. A drawback is the inability to create poles. As such resonant behavior beyond 600 Hz can only be approximated using high-order models, which weights heavy on the implementation.

The effectiveness of the method relates to the number of perturbed-parameter experiments. The FIR filter transfer matrix contains 30 filters each having $o + 1$ coefficients. Each coefficient generally associates with a perturbed and a non-perturbed parameter experiment. The number of experiments over the iterations k thus becomes: $30 \times (o + 1) \times 2 \times k$. With $o = 4$ and $k = 10$ this implies 3000 experiments!

For the scanning stage example it follows that the number of iterations can be kept small: $k = 1$. Additionally the experiments without perturbation are conducted at $\theta = 0$. The latter reduces the number of experiments almost by a factor of two: one experiment remains at $\theta = 0$. The largest reduction, however, stems from the FIR filter structure. Only one perturbation experiment per FIR filter is required irrespective the filter order o . This is because the error responses related to either the n -th or the first perturbed coefficient are equal apart from an $n - 1$ time delay [16]. As a result 31 perturbed-parameter experiments are conducted.

Under optimized decoupling the cross-talk error signal behavior is significantly improved. Less clear (at least in the experiments) are the improvements obtained in frequency-domain decoupling. Reducing the cross-talk error signals does not necessarily imply improving the degree of decoupling. The optimization incorporates all cross-talk error effects, whether decoupling-related or not. This explains why in simulation, which neither involves disturbances nor model uncertainties, major improvements are obtained in frequency-domain decoupling that clearly hint toward improved conditions for SISO control.

ACKNOWLEDGMENTS

The first author likes to acknowledge the useful comments received from Prof. Jan van Eijk during the finalization of this paper.

REFERENCES

- [1] Mishra S, Yeh W, and Tomizuka M. Iterative Learning Control Design for Synchronization of Wafer and Reticle Stages. Proc. American Control Conference, Seattle, Washington. 2008; 3908-3913.
- [2] Holmes M, Hocken R, and Trumper D. The Long-Range Scanning Stage: a Novel Platform for Scanned-Probe Microscopy. Precision Engineering. 2000; 24: 191-209.
- [3] Ho Y-C. On Centralized Optimal Control. IEEE Transactions on Automatic Control. 2005; 50(4): 537-538.
- [4] Wal van de MMJ, Baars van G, Sperling F, and Bosgra O. Multi-Variable $\mathcal{H}_\infty/\mu$ Feedback Control Design for High-Precision Wafer Stage Motion. Control Engineering Practice. 2002; 10: 739-755.
- [5] Vaes D, Swevers J, and Sas P. Optimal Decoupling for MIMO-Controller Design with Robust Performance. Proc. American Control Conference, Boston, Massachusetts. 2004; 4601-4606.
- [6] Xu K-J, and Li C. Dynamic Decoupling and Compensating Methods of Multi-Axis Force Sensors. Trans. Instrumentation and Measurement. 2000; 49(5): 935-941.
- [7] Meulen SH van der, Tousain RL, and Bosgra OH. Fixed Structure Feedforward Controller Design Exploiting Iterative Trials: Application to a Wafer Stage and a Desktop Printer. J. Dynamic Systems, Measurement, and Control. 2008; 130(5):051006.1-051006.16.
- [8] Heertjes MF, Hennekens D, Van Engelen A, Steinbuch M. Dynamic Decoupling in Motion Systems using a Gradient Approximation-Based Algorithm. Proc. Conference on Decision and Control, Shanghai, P.R.China. 2009; 5086-5091.
- [9] Vaes D, Swevers J, and Sas P. Experimental Validation of Different MIMO-Feedback Controller Design Methods. Control Engineering Practice. 2005; 13: 1439-1451.
- [10] Gunnarsson S, Collignon V, and Rousseaux O. Tuning of a Decoupling Controller for a 2×2 System using Iterative Feedback Tuning. Control Engineering Practice. 2003; 11(9): 1035-1041.
- [11] Hjalmarsson H, Gunnarsson S, and Gevers M. A Convergent Iterative Restricted Complexity Control Design Scheme. Proc. Conference on Decision and Control, Lake Buena Vista, Florida. 1994; 1735-1740.
- [12] Spall JC. (1997) A One-Measurement Form of Simultaneous Perturbation Stochastic Approximation. Automatica. 1997; 33(1): 109-112.
- [13] Spall JC. Feedback and Weighting Mechanisms for Improving Jacobian Estimates in the Adaptive Simultaneous Perturbation Algorithm. IEEE Trans. Automatic Control. 2009; 54(6): 1216-1229.
- [14] Ruszczyński A. Nonlinear Optimization.. Princeton University Press, New Jersey: 2006.
- [15] Eckhard D, and Bazanella AS. Optimizing the Convergence of Data-Based Controller Tuning. Proc. European Control Conference, Budapest, Hungary. 2009; 910-915.
- [16] Baggen M, Heertjes MF, and Kamidi R. Data-Based Feed-Forward Control in MIMO Motion Systems. Proc. American Control Conference, Seattle, Washington. 2008; 3011-3016.
- [17] Skogestad S, and Postlethwaite I. Multivariable Feedback Control; Analysis & Design. Wiley, New York: 2005.
- [18] Huusom JK, Poulsen NK, and Jørgensen SB. Improving Convergence of Iterative Feedback Tuning. J. Process Control. 2009; 19(4): 570-578.

POSITION CONTROL AND ACTIVE DAMPING

Christopher Mock
R&D Electronics

Physik Instrumente (PI) GmbH & Co. KG ·
Auf der Römerstr. 1 · 76228 Karlsruhe, Germany

INTRODUCTION

Miniaturization is a well known keyword in terms of scientific and technical progress. In regard to electronics and mechanical components, it is subject to the so called mechatronics. This leads to constant increasing requirements regarding tools and measuring equipment. To qualify, for example, structures below 13 nm, the measuring device must provide a resolution far below one nanometer.

Sub-nanometer resolution can be achieved by piezo actuator systems. The effect used by the well-known piezoelectric lighters-voltage generation when pressure is applied-is reversible. The "inverse piezo effect" is to convert electrical energy to mechanical energy. In combination with a preloaded mechanical spring system, a displacement is the consequence of a voltage change. Since this effect is affected by great nonlinearity, a closed loop is needed for a positioning system which is based on piezo actuators. Developing a positioning system with sub-nanometer resolution is of course a challenge but will not be discussed here. It is presumed that positioning system and measuring device provide the required resolution.

In the following, the requirements concerning dynamic and static behavior of such a system are discussed. In particular with scanning applications, accurate and constant velocity and small tracking errors are essential.

REQUIREMENTS

A piezo actuator system in the form of preloaded mechanical spring system can be regarded as a one-mass-spring system.

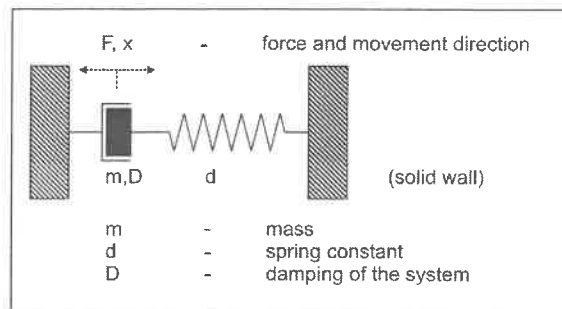


FIGURE 1 One-mass-spring system

According to experience, piezo actuator systems are damped inadequately. For that reason, the resonance of the system cannot be neglected with closed loop systems. Methods have to be found to avoid amplified movement at resonant frequency. Due to the inadequate damping it is of particular importance to consider distortion forces that are not excited by the piezo element.

It is possible to avoid exciting resonance by decreasing the closed-loop speed. But due to the dynamic requirements this is not tolerable.

The requirements concerning a closed-loop piezo positioning system are as follows:

- High closed-loop bandwidth
- Low phase shift between target and real position
- High shape accuracy
- High distortion rejection (especially for the resonant frequency and higher frequencies)
- High curve constancy regarding the target signal
- Small tracking error

As mentioned above, the relation of position displacement to applied voltage is nonlinear which is mainly due to a hysteresis effect.

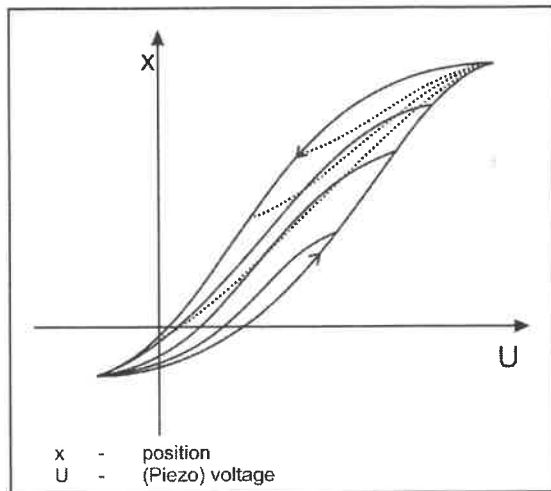


FIGURE 2 Piezo hysteresis

For open-loop systems the actual position depends on the past position and the gradient of voltage. The past position was dependent on the last preceding position and again on the voltage gradient. To describe a piezo hysteresis, hence a large memory would be needed to rebuild the sequences of past positions and voltage gradients. Thus for open-loop systems the position is never clear. In addition to this hysteresis effect, the position is affected by drift. To compensate these nonlinear effects it is therefore essential to close the loop using a position sensor signal. By closing the loop the system will be linearized statically. But in dynamic applications nonlinear effects will still occur. When a triangle target signal is applied, for example, the system will not produce a constant tracking error as expected for linear systems. The tracking error will change over time while the target signal has a constant slope. This behaviour results from the fact that the piezo systems velocity is changing while it should be constant. This is an unwanted effect for all applications assuming the piezo actuator system to be linear.

Thus a piezo actuator system can be described and classified by the following criteria:

- Relation of position displacement to applied voltage
- Resonant frequency

APPROACH

Since closed-loop bandwidth is limited, high-order nonlinearity has to be compensated by a method. The control method has to handle the

resulting effects of the inadequate damped system and the relatively slow drift effects of the piezo actuator.

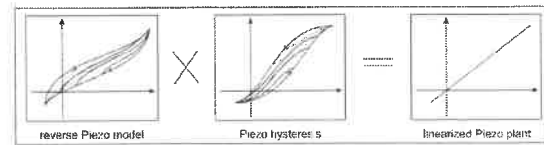


FIGURE 3 Hysteresis compensation

Assuming nonlinear effects are now reduced to a minimum, the plant of a piezo actuator system can be regarded as an inadequate third order system consisting of the inadequately damped part and the entirety of all well-damped parts like amplifier bandwidth, delays and sensor bandwidth.

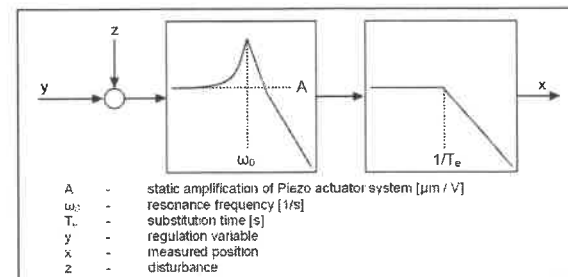


FIGURE 4 model of the linearized piezo actuator system

The transfer function can be reduced to the following La Place description.

$$F(s) = \frac{A}{1 + \frac{2 \cdot D}{\omega_0} \cdot s + \frac{1}{\omega_0^2} \cdot s^2} \cdot \frac{1}{1 + T_e \cdot s}$$

FIGURE 5 Transfer function of a piezo actuator system

One method to avoid excitation of the resonance is to add a notch filter to the regulation variable y. The disadvantage is that the distortion is still exciting the resonance, and by adding a notch filter exactly this distortion can not be compensated. When analyzing the pole-null-diagram of the model a better method becomes apparent.

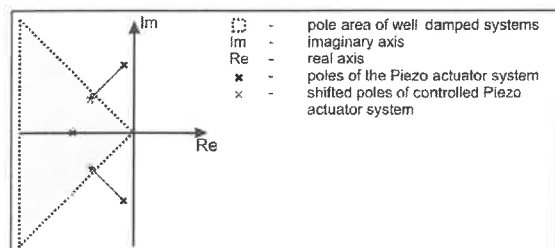


FIGURE 6 Pole-null-diagram of a piezo actuator system

The ideal way is to shift the poles of the piezo actuator system into the well-damped pole area. Thus the amplification at the resonant frequency will be damped actively without elimination of any frequency components regarding the signals regulation variable y or the sensor signal x .

SOLUTION

Hysteresis compensation:

The company PI has a patent for a reverse piezo model (reference number DE102004019052A1). This model has been expanded with a polynomial to inhibit higher order nonlinearities that cannot be linearized by the reverse piezo model.

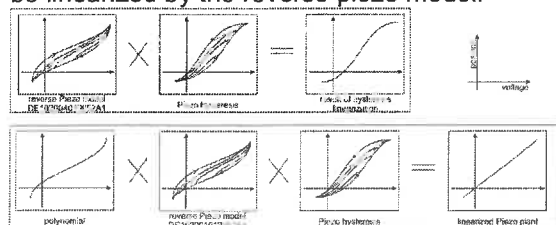


FIGURE 7 Expanded piezo linearization

FIGURE 7 shows the behavior of the reverse piezo model in combination with the piezo actuator system. The result is a monotonic relation between voltage and resulting position. It still shows a higher order error that can be linearized by the reversed polynomial. After applying the polynomial the relation between voltage and resulting position is constant ($A=\text{const.}$).

Resonant frequency in closed-loop operation:

In the *APPROACH* section, the reduced model of a piezo actuator system was shown. The transfer function given in FIGURE 10 can be represented by the following block diagram.

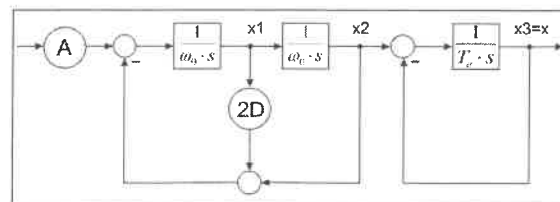


FIGURE 8 Piezo actuator system block diagram

To normalize the transfer behavior to a static amplification of 1 the regulation variable y will be divided by the quotient A . Thus this factor could be neglected for further considerations. Now being able to feed back the weighted intermediate states $x1$, $x2$ and $x3$ to the system, closed-loop dynamics can be adjusted by the so called k -coefficients.

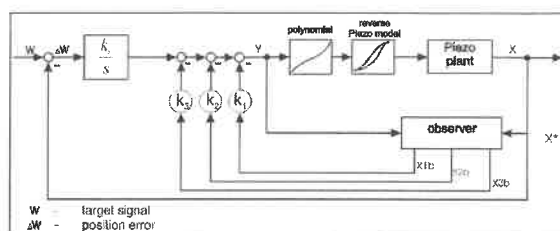


FIGURE 9 Control structure

The overall control is done by a simple integral action controller to eliminate static and drift distortions and to be able to reach the exact target position. An observer was built to generate the state signals $x1$, $x2$ and $x3$ (observed signals are called $x1b$, $x2b$ and $x3b$). The La Place closed-loop transfer function can be described by the following equation:

$$F_o(s) = \frac{k_i \cdot s^{-1}}{k_i \cdot s^{-1} + (1 + k_1) + \left(\frac{2 \cdot D}{\omega_0} + T_e + k_2 \right) \cdot s + \left(\frac{2 \cdot D}{\omega_0} \cdot T_e + \frac{1}{\omega_0^2} + k_3 \right) \cdot s^2 + \left(\frac{1}{\omega_0^2} \cdot T_e \right) \cdot s^3}$$

FIGURE 10 Closed-loop transfer function

The transfer function shows that all poles of the closed-loop system can be adjusted by the k -coefficients. Thus a well-damped system can be created without any signal-eliminating filters. The main advantage is that no delays are added to the system. Hence the phase shift between target and real position signal can be minimized. The excitation of the resonance is damped actively even for distortion signals. By choosing an ideal combination of the k -coefficients (e.g. method of the double relations [1] $DV=2$), high frequencies (in relation to the closed-loop bandwidth) are damped by the structure with up to -80 dB per decade.

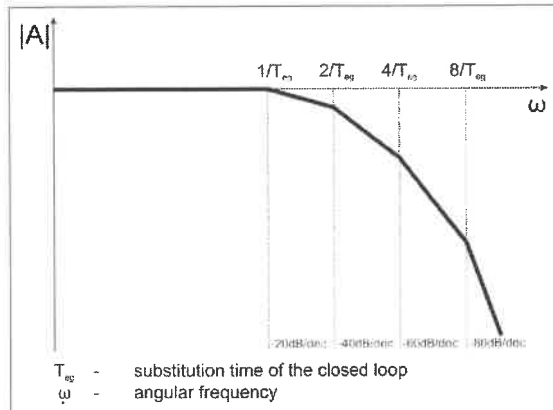


FIGURE 11 Structural damping of high frequencies

RESULTS OF MEASUREMENTS

Hysteresis compensation:

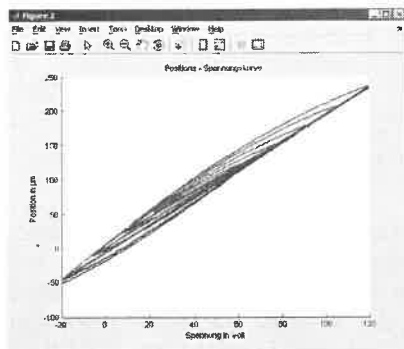


FIGURE 12 Real hysteresis

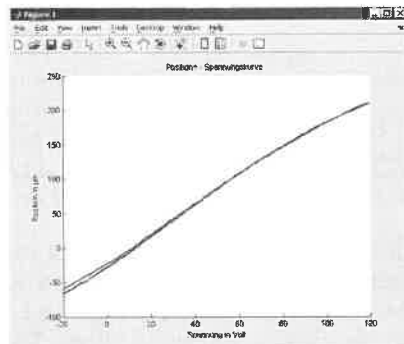


FIGURE 13 With reverse piezo model

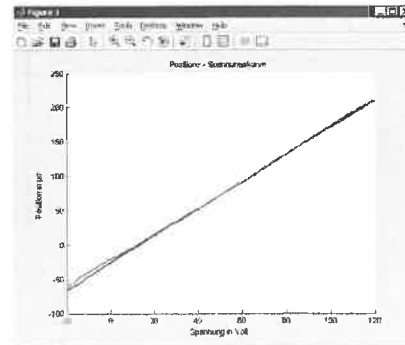


FIGURE 14 With reverse piezo model and polynomial

FIGURE 12 to FIGURE 14 show that the piezo actuator system can be linearized. The result is a nearly constant relation between applied voltage and position. The effect on closed-loop operation is shown below by figures comparing standard PID control with notch filter and the so called advanced piezo control described in this document. The target was a 100 μm triangle signal with 5 Hz.

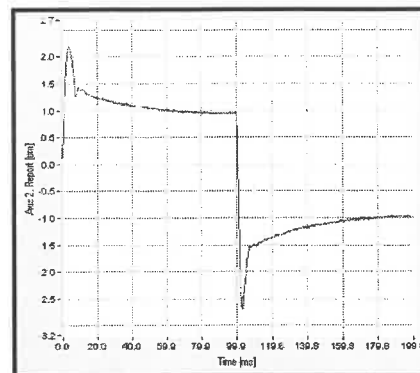


FIGURE 15 Tracking error with PID

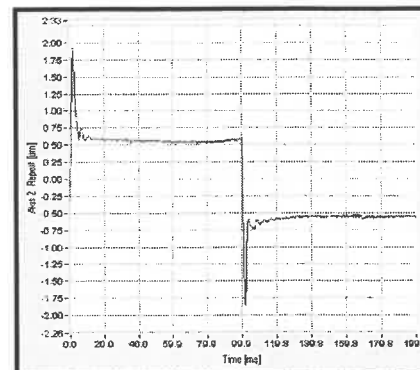


FIGURE 16 Tracking error with advanced piezo control

The tracking error is reduced to 50%. The position error is constant and the system behaves in the same manner for positive and negative slopes.

Resonant frequency in closed-loop operation:

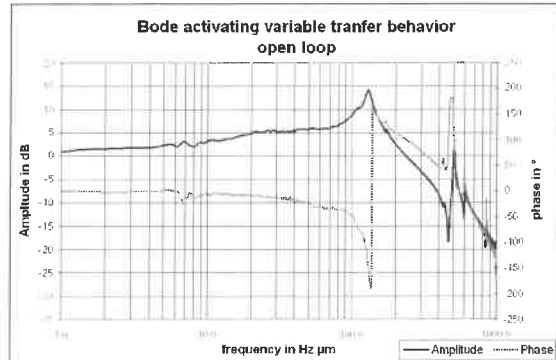


FIGURE 17 Open-loop bode diagram of a piezo actuator system

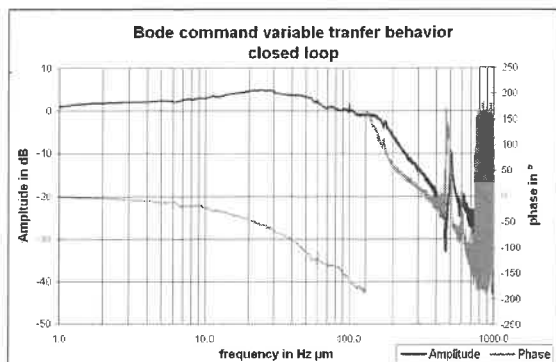


FIGURE 18 Closed-loop bode diagram with advanced piezo control

The resonance is not excited. This is achieved without eliminating any frequency component from the target signal w or the regulation variable y .

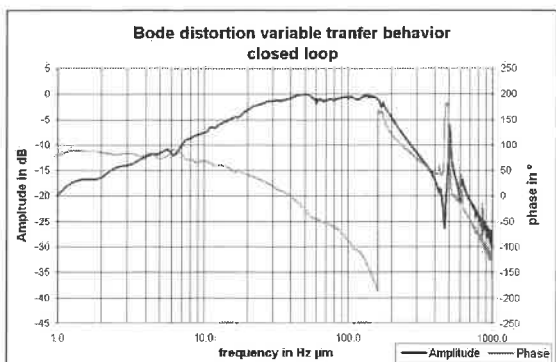


FIGURE 19 Closed-loop distortion bode diagram $x(s) / z(s)$

The distortion signal z is not exciting the resonance. Low frequency parts are damped.

CONCLUSION

The control structure provides not only position control but also damps the system resonance actively. Thus both position control and active damping are provided.

REFERENCES

- [1] Prof. Franz Kolb papers of lecture "closed loop control II".
- [2] Prof. Franz Kolb papers of lecture "synthesis of closed loop control".
- [3] Prof. Franz Kolb papers of lecture "digital closed loop control".
- [4] Prof. Franz Kolb papers of lecture "state space methods".
- [5] Heinz Unbehauen: Regelungstechnik III, Vieweg, 2000, ISBN 3528533498.
- [6] Otto Föllinger: Regelungstechnik. Hüthig Verlag, ISBN 3-7785-2336-8.

CONTROLLER DESIGN FOR PRECISION X-Y STAGE USING SENSOR/ACTUATOR AVERAGING AND ACTIVE DAMPING

Dan Gordon¹, Kaan Erkorkmaz¹

¹Precision Controls Laboratory, Dept. of Mechanical & Mechatronics Engineering
University of Waterloo, 200 University Ave. West, Waterloo, ON, N2L 3G1

INTRODUCTION

The demand on the manufacturing industry to produce miniaturized and micromachined components continues to rise, particularly for biomedical, automotive, and opto-mechatronic applications. Traditional machine tools no longer present a viable means of producing such parts, due to their relatively low accuracy, as well as excessive foot print and power consumption. To meet these manufacturing needs, a new generation of micro/meso machine tools has started to emerge. At the University of Waterloo, a prototype x-y stage has been developed as a new tool/part positioning means for such machine tools.

The stage is designed around the T-type gantry concept, with a single guideway and a secondary linear motor to achieve active stiffness and damping enhancement. The coupled nature of the dual-drive gantry presents a major controls challenge, which has been addressed using sensor/actuator averaging and active vibration damping. Stability analyses have been carried out using the multivariable control framework. The details of the controller design are presented in the rest of this paper.

EXPERIMENTAL SETUP

The experimental setup is a T-type gantry and worktable configuration, as shown in FIGURE 1. The gantry (125 kg) is supported on a precision granite surface by four Ø80 mm porous air bearings. It is constrained to move along the y axis by four porous 150x75 mm² air bearings that are preloaded against a precision granite guideway. These bearings provide both lateral and yaw stiffness for the stage. The worktable (33 kg) is supported by a 12"x12" Vacuum Pre-Loaded (VPL) air bearing on the granite base and is guided along the stainless steel gantry beam by four preloaded air bearings (100x50 mm²). The VPL decouples the worktable and gantry from vertical loads and also increases the vertical and pitch stiffness of the table. Actuation of the gantry is achieved through two linear motors. The primary motor is located next to the

guideway, while the secondary motor is located at the opposite end to actively stiffen the beam. The worktable is actuated by a single linear motor secured to the gantry. The stage has an overall work area of 300x300 mm² and is capable of 1 g acceleration to a maximum speed of 1 m/s. High resolution linear encoders provide feedback with a sinusoidal period of 4 µm, which can be sub-interpolated by 4096 times.

CONTROLLER DESIGN

During identification tests, it was found that the x axis can be considered independent of the gantry excitation. The open loop Frequency Response Function (FRF) was measured, and was found to be flat up to 210 Hz. The control law was developed using the loop shaping technique, which achieved a 140 Hz crossover frequency with a phase margin of 25 degrees and peak sensitivity of 2.90. In addition, inertia compensation and a trajectory prefilter were added in feedforward to improve command tracking. The trajectory prefilter is designed to remove the correlations of velocity, acceleration, jerk and snap (4th derivative) commands from the tracking error, and is tuned using Least Squares using the results of a single tracking experiment [1].

Due to the strong coupling between the primary and secondary actuators through the gantry beam dynamics, the y axis controller design proved to be more challenging. As a result, a multivariable control framework was used to avoid serious stability and bandwidth limitations.

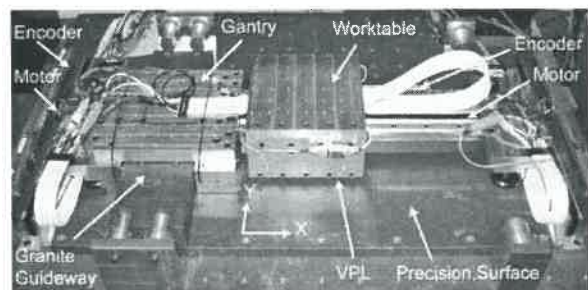


FIGURE 1. Experimental setup with T-type gantry and worktable configuration.

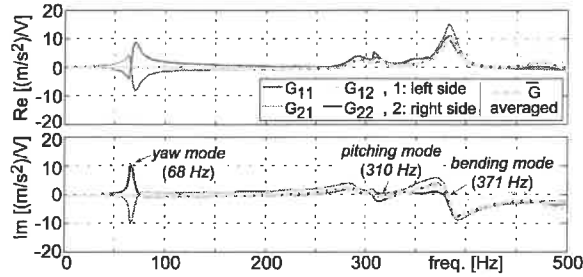


FIGURE 2. Open loop acceleration FRF's y axis

FIGURE 2 shows the experimentally recorded direct and cross FRF's for the gantry. The first mode at 68 Hz is most prevalent. This mode originates from the rigid body rotation of the gantry constrained by a torsional spring-like effect of the four air bearings preloaded against the granite guideway. The second mode, at 310 Hz, is attributed to the pitching of the gantry. These two modes do not change much with x axis position. The third mode at 371 Hz corresponds to the bending of the beam, and is more strongly dependent on the table position. When the table is moved from the extreme left position ($x = -150$ mm) to the extreme right ($x = +150$ mm), this mode shifts from 365 Hz to 380 Hz.

The table position also affects the mass distribution reflected on the drives. Defining m_1 and m_2 as the control signal normalized mass values for the left and right hand sides of the gantry, the mass distribution ratios are defined as:

$$\alpha_1(x) = m_1(x)/m, \alpha_2(x) = m_2(x)/m \quad (1)$$

Where $m = m_1 + m_2$ is a constant. The values of m_1 and m_2 are determined by correlating the current demand for the motors with the commanded acceleration and velocity profiles using data logged from tracking experiments. Three experiments are performed, with the x axis at the left (-150 mm), centre (0 mm) and right (+150 mm). For our setup, the mass distribution changed by 8.68 percent throughout the range of the table motion, with the primary motor actuating 51.2 percent of the mass when the table was in the centre position.

The developed controller is shown in FIGURE 3. Position measurements $[y_1, y_2]$ from the encoders are utilized in a state feedback

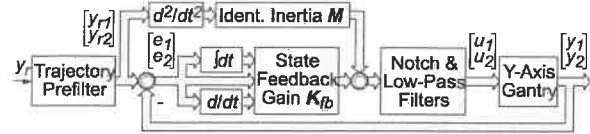


FIGURE 3. Multivariable control structure developed for the y axis

scheme. The state vector comprises of the tracking errors and their time integrals and derivatives: $x' = [\int e_1 dt, \int e_2 dt, e_1, e_2, \dot{e}_1, \dot{e}_2]$. The feedback gain has the following structure:

$$K_{fb}(x) = [K_i(x) : K_p(x) : K_d(x)]_{2 \times 6} \quad (2)$$

The first step to compute K_{fb} was to design a PD controller in the form $k_p + k_d s$, considering the gantry as a rigid body. For a desired natural frequency ω_n and damping ratio, ζ :

$$k_p = m\omega_n^2, \quad k_d = 2\zeta\omega_n m \quad (3)$$

The values were chosen as $\omega_n = 45$ Hz and $\zeta = 1.2$. This controller uses the centre of mass (CoM) position estimated from the two encoders: $\bar{y} = \alpha_1 y_1 + \alpha_2 y_2$. Its output ($\bar{u}$) is distributed to the motors proportionally to the mass they actuate: $u_1 = \alpha_1 \bar{u}, u_2 = \alpha_2 \bar{u}$. Hence, K_p and K_d become:

$$K_p(x) = k_p \begin{bmatrix} \alpha_1^2 & \alpha_1 \alpha_2 \\ \alpha_2 \alpha_1 & \alpha_2^2 \end{bmatrix}, \quad (4)$$

$$K_d(x) = k_d \begin{bmatrix} \alpha_1^2 & \alpha_1 \alpha_2 \\ \alpha_2 \alpha_1 & \alpha_2^2 \end{bmatrix}$$

As seen in FIGURE 2, the application of this scheme nearly removes the yaw mode from the averaged transfer function being controlled ($\bar{G} = \alpha_1^2 G_{11} + \alpha_1 \alpha_2 (G_{12} + G_{21}) + \alpha_2^2 G_{22}$), while avoiding phase lag that is normally associated with notch filtering a vibration mode. This allows the control performance to be significantly improved, by reducing the negative effect of the yaw mode on closed loop stability, similar to the methodology in [2].

K_i is designed to eliminate steady state errors at the servo level rather than the CoM. This is because opposing servo errors may give the

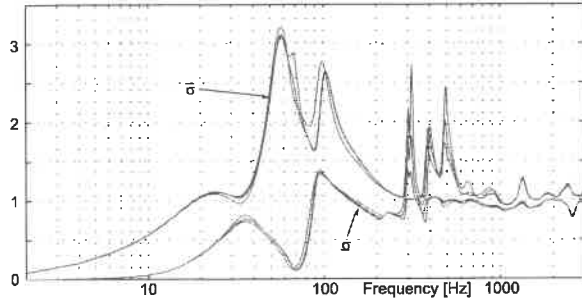


FIGURE 4. Sensitivity singular values for y axis controller with x axis at left, centre and right

false impression that the CoM tracking error is zero. Hence, it is computed as:

$$K_i(x) = k_i \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \quad (5)$$

Above, $k_i < k_p k_d / m$ is required for stability based on Routh-Hurwitz criterion; assuming rigid body dynamics. The practical limit for k_i was actually lower due to machine base vibrations.

In addition to the derivative action for CoM control, active damping is also injected to attenuate yaw vibrations which may be excited by external disturbances. This is crucial for enhancing the stability margins, which enables higher control bandwidth and dynamic stiffness to be achieved. The active damping is designed to emulate a torsional damper, which increases the yaw damping by a factor of 32x. The output is distributed to the drives based on their mass ratios. Hence, the final form of K_d becomes:

$$K_d(x) = k_d \underbrace{\begin{bmatrix} \alpha_1^2 & \alpha_1 \alpha_2 \\ \alpha_2 \alpha_1 & \alpha_2^2 \end{bmatrix}}_{\text{Derivative action for CoM control}} + \underbrace{c \begin{bmatrix} \alpha_1 & -\alpha_1 \\ -\alpha_2 & \alpha_2 \end{bmatrix}}_{\text{Active damping to attenuate yaw vibrations}} \quad (6)$$

In practical implementation, additional filters were also required to stabilize the closed loop and avoid poor robustness. These filters were designed by applying multivariable stability and sensitivity analyses. Considering the experimentally recorded open loop frequency response matrix $G(j\omega)$, and the feedback controller frequency response $K(j\omega)$ computed analytically, the multivariable loop transfer matrix is constructed as $L(j\omega) = G(j\omega)K(j\omega)$. Stability is guaranteed when the locus of $\det\{I + L(j\omega)\}$

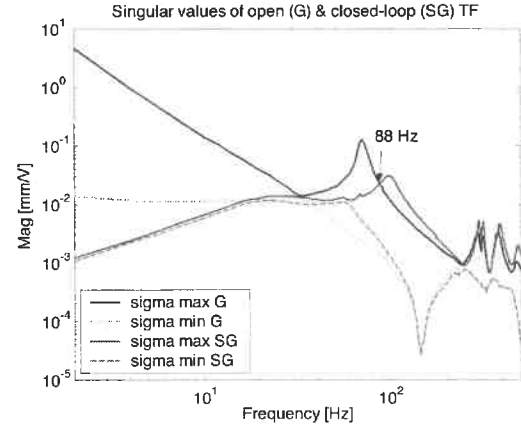


FIGURE 5. MIMO singular values for open and closed loop disturbance responses

makes no counter-clockwise encirclement of the origin in the complex plane [3]. Also, the maximum singular value of the sensitivity function ($\bar{\sigma}(S(j\omega)) = \bar{\sigma}\{(I + L(j\omega))^{-1}\}$), shown in FIGURE 4, was inspected as a margin of robust stability. This is useful in determining critical frequencies that need to be low pass or notch filtered. When the feedback design was complete, the cross over frequency was 98 Hz, and the peak sensitivity 3.12. Comparing the open and closed loop disturbance responses (FIGURE 5), the frequency up to which feedback provided an improvement was 88 Hz.

To enhance command tracking, inertial forces were compensated and a trajectory prefilter was added, as shown in FIGURE 3. The prefilter was tuned by conducting five tracking experiments at $x = -150, -75, 0, +75, +150$ mm. The individual prefilter terms were then gain-scheduled based on the commanded x axis position.

EXPERIMENTAL RESULTS

The servo performance for the individual axes at different speeds is shown in FIGURE 6. The displacements were 1 mm and 200 mm in the 1 mm/s and 200 mm/s tests. At 200 mm/s, the x and y axes maintain accuracies of 1.71 μm and 1.87 μm . The maximum tracking errors at 1 mm/s are around 100 nm. The holding accuracy is better than ± 80 nm in both axes.

Individual axis tracking errors and contouring error for a $\varnothing 1$ mm circular trajectory are shown in FIGURE 7. The test was conducted at 1 mm/s feed, with 2.6 m/s^2 acceleration and 10 mm/s^3 jerk along the feed direction. The servo errors are less than 91 nm for the x axis and primary y

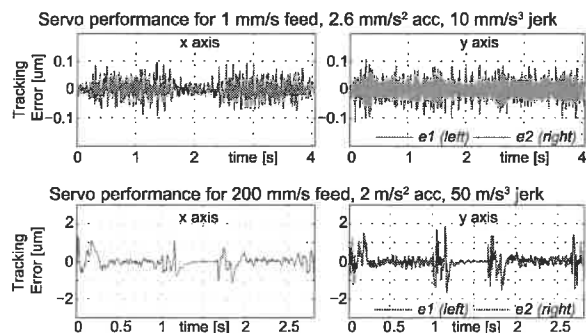


FIGURE 6. Servo errors at different speeds.

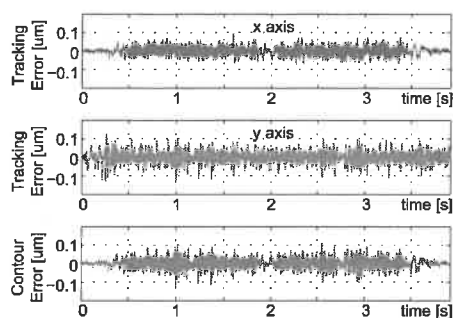


FIGURE 7. Tracking and contour errors for a circular trajectory.

axis, and 121 nm in the secondary y axis. The maximum contour error is 73 nm when evaluated from the primary side of the y axis, and 111 nm from the secondary. FIGURE 8 shows corner tracking during a diamond contouring test with the same feed, acceleration and jerk values. The maximum error during diamond contouring is around 110 nm.

In FIGURE 9, the contribution of active damping

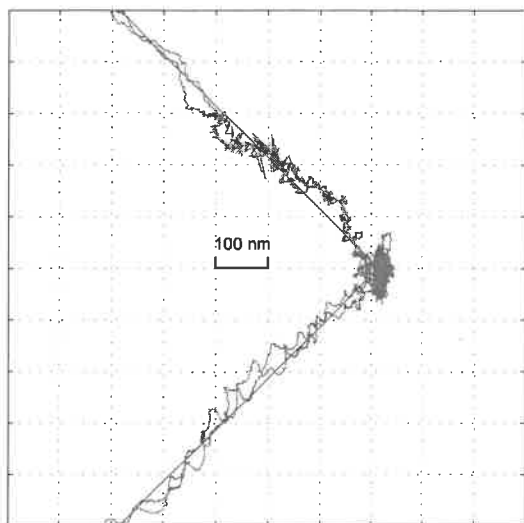


FIGURE 8. Corner tracking during diamond contouring.

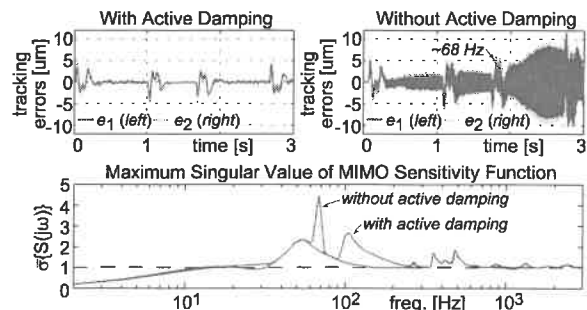


FIGURE 9. Contribution of active vibration damping on y axis (gantry) control.

is shown. When active damping is removed, the gantry goes into yaw vibrations at 68 Hz. This is confirmed by the high sensitivity ($\bar{\sigma}(S(j\omega))$) around this frequency. When active damping is enabled, the sensitivity peak reduces from 4.46 to 2.65 and shifts to a higher frequency of 104 Hz. The tracking tests were performed at 200 mm/s. In order to keep stability in the absence of active damping, the controller was detuned resulting in larger tracking errors. It is noteworthy that when controlling the y axis with independent PID controllers, the achievable crossover frequency was only 16.4 Hz and the active disturbance rejection band was 8.2 Hz. The corresponding values in the presented multivariable design are an order of magnitude higher.

ACKNOWLEDGEMENTS

This research is sponsored by CFI, AUTO21 and Ontario MRI granting agencies in Canada. Industrial contributions were provided by New Way Precision, Heidenhain, ETTEL, and Rock of Ages. The authors gratefully acknowledge the help of CNC technicians Mr. Robert Wagner and Mr. Jason Benninger in building the setup.

REFERENCES

- [1] Gordon, D.J., Erkorkmaz K., Accurate Control of Ball Screw Drives using Pole-Placement Based Vibration Damping and Optimized Trajectory Pre-Filtering, Proc. ASPE, Monterey, CA. 2009; 211-214.
- [2] Weng, M C, Lu X, Trumper D L. Vibration Control of Flexible Beams using Sensor Averaging and Actuator Averaging Methods. IEEE Trans. on Control Systems Technology. 2002; 10/4:568-577.
- [3] Skogestad S, Postlethwaite I. Multivariable Feedback Control – Analysis and Design (2nd Edition). Wiley. 2005.

FAST LONG RANGE ACTUATOR (FLORA II) FOR FREEFORM OPTICAL SURFACES

Thomas A. Dow, Kenneth Garrard, Alexander Sohn
Precision Engineering Center
North Carolina State University
Raleigh, NC, USA

INTRODUCTION

This paper describes the design of a new tool for machining Non-Rotationally Symmetric (NRS) surfaces. NRS optical surfaces have traditionally been machined using slow-slide servo techniques using the massive machine axes and low spindle speeds or limited range servos driven by piezoelectric actuators. A Fast Long Range Actuator (FLORA I) was constructed in 2005 with a goal of machining NRS surfaces with a sag up to 4 mm at 20 Hz with a form error of less than 150 nm and a surface finish of 5 nm RMS. This is a challenging undertaking because the range is 4 mm, the resolution 1 nm and at the 250 mm/sec maximum tool speed, the tool will move 12 μm in the controller update cycle (50 μsec).

The new FLORA II system uses a hollow, triangular SiC piston that is half the mass of the original aluminum honeycomb, porous carbon bearings have replaced the orifice design and a voice coil motor is in place of 3-phase linear motor. These changes have made a smaller lighter more controllable design. Table 1 compares the two designs indicating a substantial reduction in package mass and physical size [1,2].

Table 1 Comparison of Flora I and II

	Flora I	Flora II
Total Mass, Kg	36	7
Piston Mass, Kg	0.65	0.36
End-Load Stiffness, N/ μm	53	8
First Natural Freq, Hz	1000	2700
Housing Height, mm	130	90
Housing Width, mm	205	110
Housing Length, mm	305	208

DESIGN DETAILS

The design of the structure and the integration of the porous bearings was a demanding task. For example, the air bearings must be modeled and coupled with the piston and the structure. If the piston had excessive mass or the bearing

possessed inadequate stiffness, the resulting low natural frequencies could decrease performance. The thickness of the air film was a critical requirement of the new design with a goal of 5 μm . This air film is created through the expansion of the housing due to the air pressure as well as the compression of piston. The influence of each component was evaluated to create the shape and stiffness of the housing.

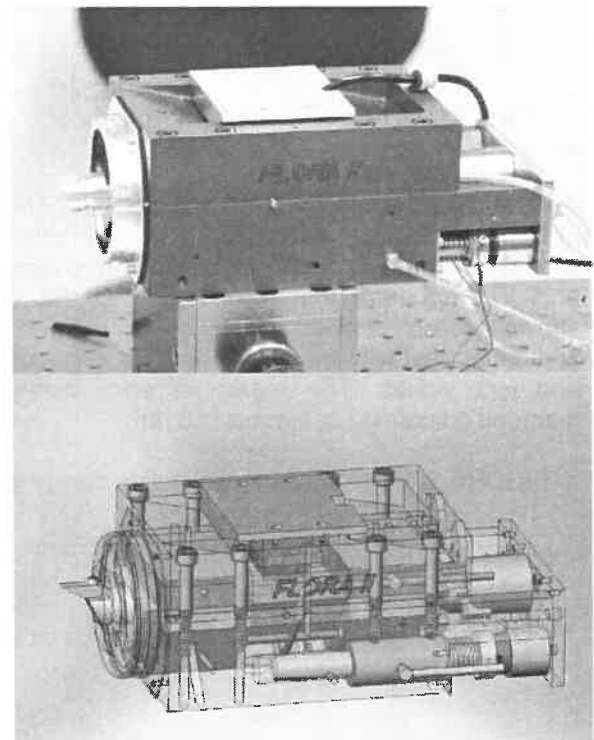


FIGURE 1. Flora II configuration, photograph at top and transparent design view at bottom.

Figure 1 shows a photograph (top) and a transparent image (bottom) of the final design of Flora II. The tool holder is on the left, encoder access is at the top, the main voice coil motor is at the right above the two counterbalance motors. The transparent view of the design shows further detail such as the silicon carbide piston, custom New Way air bearings,

counterbalance system, and air supply plumbing.

BEI Kimco voice coil motors (VCMs) were selected to actuate the main piston as well as the counter balances. The single-phase characteristics of the VCMs were advantageous when compared to the linear three phase brush motor implemented in FLORA I. Two Trust Automation linear amplifiers were chosen to power the VCMs based on their low noise and high bandwidth. Position feedback was provided by a Sony LASERSCALE encoder with sub-nm resolution and a maximum velocity over 250 mm/sec. A PC with PMDi D100 controller was used for motion commands and to create a user interface to manage the different modes of operation of the FLORA II. The major features of the design were:

1. A SiC piston with hollow, triangular cross-section was selected based on several candidates shapes.
2. Replacing the orifice type air bearings with porous carbon air bearings lead to higher stiffness and less parasitic piston motion.
3. The housing, including counterbalance, was significantly smaller and lighter than its predecessor. The structure was designed to create a uniform air gap of 5 μm .

CONTROLLER DESIGN

To design the control filter, the dynamics of the system – amplifier, motor, structure, position measurement – were identified.

Open Loop System Identification

The open loop system was identified by exciting it with a series of different single frequency sine waves. For the FLORA II, it was determined that a range of 0 to 4 kHz in 10 Hz steps would be sufficient to define the dynamics.

Once the data for amplitude ratio and phase was collected and analyzed, it was plotted as a function of frequency (rad/sec) in 2. This transfer function relates the output in mm to an input in volts to the amplifier. The data on the top in Figure 2 shows an amplitude reduction of 40 dB per decade and a phase lag of about 180° at frequencies greater than 10 Hz (60 rad/sec). A model was fit to the experimental transfer function shown on the bottom of Figure 2 by combining expressions that describe the behavior of the different components in the system. Because the input command is voltage (which produces current or force in the motor

and acceleration in the piston) and the output is position, there will be a theoretical reduction in amplitude of 40 dB for each decade of increased frequency and 180° phase shift between input (force) and output (displacement).

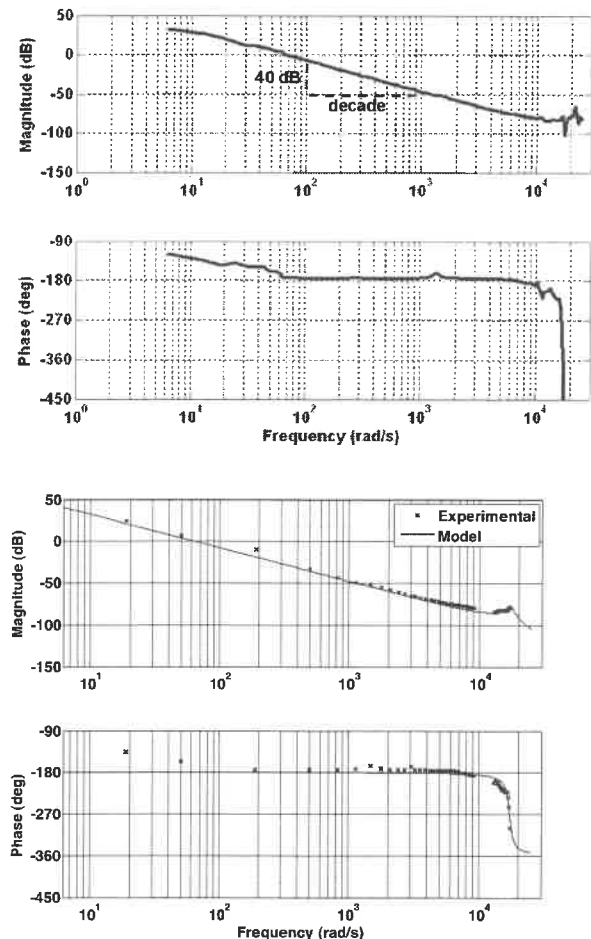


FIGURE 2. Open loop transfer function of FLORA II from sine sweep. Top is the original data and bottom shows the fit of the model.

The gain G_A of the system (magnitude at a specific frequency) is a function of the amplifier gain and the mass of the piston. In addition to the linear change with frequency, there is a peak at about 2800 Hz (17.5 rad/sec) and a 180 phase shift indicating a structural resonance. Combining these effects with values of $\zeta = 0.05$, $\omega_n = 17,216$ rad/s and $G_A = 4158$ mm/V, the following equation approximates the experimental transfer function. This matches the measured data as shown in Figure 2.

$$\frac{G_A \omega_n^2}{s^4 + 2\zeta \omega_n s^3 + \omega_n^2 s^2} = \frac{1.23e^{12}}{s^4 + 1722s^3 + 2.96e^8 s^2}$$

Closed Loop Controller Design

The control system implemented for FLORA II is a PID controller with acceleration feed forward. A block diagram for this filter is shown in Figure 3. The gains, K_P , K_I , and K_D for the controller were determined using frequency design procedure based on the desired bandwidth.

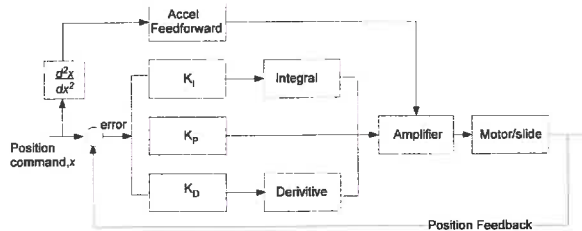


FIGURE 3. FLORA II control system

The magnitudes of the integral and derivative gains are associated in the frequency response of those two filters and the desired bandwidth, ω_b , of the system. The integral term increases the low frequency gain for steady-state accuracy while the derivative term is used to improve stability and increase response. By changing the frequency where the integral and derivative filters influence the system, the response can be modified. The standard design is to make the integral gain end at 10% of the desired bandwidth and the derivative to start at 50%. Based on these relationships, the gains are:

$$\frac{\omega_b}{10} \simeq \frac{K_I}{K_P} \quad \frac{\omega_b}{2} \simeq \frac{K_P}{K_D}$$

The shapes of three filters with different gains (described in Table 1) are shown on the right side of Figure 4. The medium and high have similar frequency characteristics but the low gain shows the effect of reduced integral gain at low frequency. Changing the gain K_P shifts the entire response curve up and down.

TABLE 2. Controller gains used for FLORA II tuning

Controller	K_P	K_I	K_D
High	200		
Medium	135	16200	0.225
Low	60	1	0.179

The right side of FIGURE 4 shows the open loop response from Figure 2 but modified by the medium gain controller. Two measures of the utility of the controller are the gain margin and the phase margin. The Phase Margin is the phase angle of the response from 180° when the system passes through 0 dB magnitude. In this

case it is about 55° which is considered good for stability. The Gain Margin is the additional gain that could be added before the system goes unstable. The 9 dB for this example is excellent. Therefore, the FLORA II system with the medium gain filter should perform well.

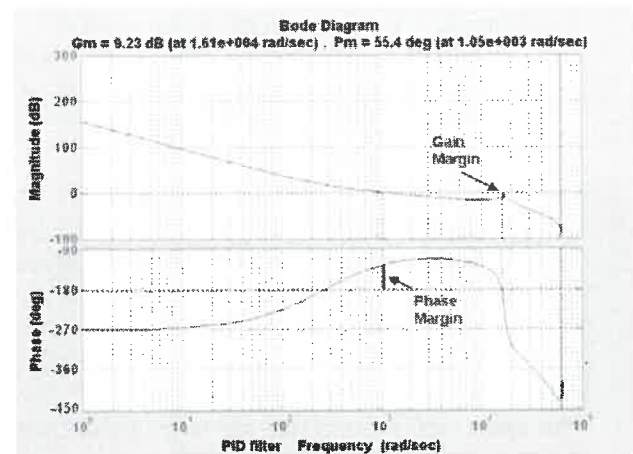
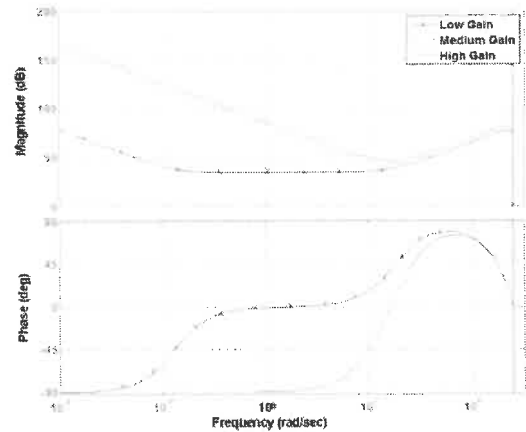


FIGURE 4. Frequency response of different PID controllers (top) and the medium gain case applied to the FLORA II dynamics (bottom).

The FLORA II controller also uses acceleration feed forward to improve the response to changes in position in addition to the control filter. The feedforward gain is determined by following error experiments and is used to adjust the phase of the command to reduce errors.

DSP Control System

The FLORA II is controlled by a PMDi DMC 100 DSP board in a PC with analog and digital I/O cards connected to the DSP with MotionWire ports [4,5]. The user interface runs on a PC and downloads DSP code to the DSP and receives

position information for the display. The highest level of the control code including is a monitor program that runs at 1 KHz and manages the high level machine operations such as checking for emergency stop conditions, posting position data on the user interface screen and allowing the user to switch among the general operating modes.

At each control cycle interrupt (50 μ sec), the DSP reads the location of each axis (radius of part or X axis position of DTM, Z axis position of the DTM, rotational position of spindle and position of the FLORA II piston from encoder) and based on the next desired position, new commands based on the acceleration and position errors, a new command is sent to the main and counter-balance motors.

OPTICAL AND FIDUCIAL FABRICATION

The ability to machine large-amplitude (> 1 mm) NRS features also makes it possible to create assembly features in the same operation as the optical surface. This is critical for freeform surfaces because the added degrees of freedom of these shapes complicate assembly.

Biconic Mirror

The M4 mirror was designed by NASA for a space-based spectrometer [5]. The mirror is an off-axis biconic toroidal surface that blends two oblate ellipses with different curvature and eccentricity in the XZ and YZ planes. This shape is used as an example of a freeform surface because a Computer Generated Hologram (CGH) is available to measure the shape created on a laser interferometer

Figure 7 shows a section of the biconic surface as a wire frame and the aperture of the mirror blank as a solid. This optic was originally machined at the PEC [3] in 2002 using a piezoelectric actuator with mechanical amplification for a 400 μ m range. Because this range was less than the sag of the M4 surface, the optical blank was machined off-axis and tilted 35.3° with respect to the fabricating axis. The actuator was also tilted with respect to the spindle axis. These two changes allowed the surface to be machined with the range available but complicated the setup of the machining operation.

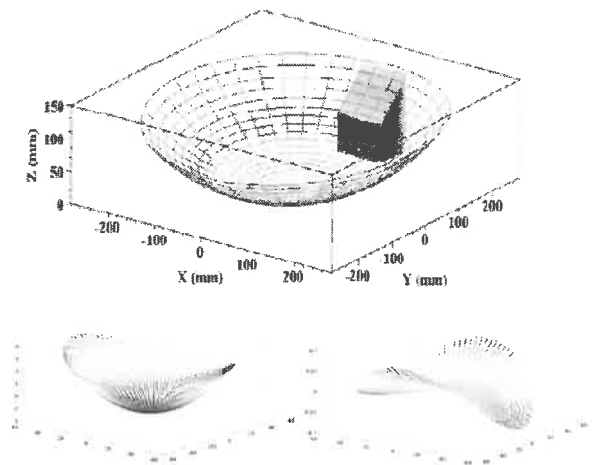


FIGURE 5. Biconic optic M4 showing (top) parent optic, (Lower-left) on-axis M4 surface and (lower-right) NRS component.

With the extended range of the FLORA II, it is possible to machine the M4 surface on axis and simplify fabrication. The lower images in FIGURE 5 show the on-axis M4 surface with a maximum aperture 98 mm, the surface decentered by -227.41 mm in Y and 2.01 mm in X and rotated -35.3° around X axis. The total sag of the on-axis surface is 4.776 mm, but after removing the best sphere with a radius of 376.00 mm, the excursion of NRS component is 0.153 mm.

Fiducial Features

An example of an ideal assembly feature is the well-known kinematic coupling consisting of 3 balls on the optic that fit into three grooves on the housing. This concept could replace the time consuming shim-and-measure technique that has plagued the application of freeform surfaces in multi-element optical systems.

The shape of the mating features is a compromise between the needs for fabrication and operation of the interface. There are two main concerns.

The stress at the contact should not exceed the yield stress of the material. Since the materials are 6061-T651 aluminum, large radii at the contact point are important.

The slope of contact surfaces must be sufficient to overcome friction at the interface and allow the protruding features to fall to the bottom of the grooves.

Figure 6 shows detailed shape of the feature design. They have constant 4 mm width in the radial direction and a full cosine wave in the circumferential direction with a wave length of 8 mm. The high-point in the radial direction is at a radius of 61.5 mm. The sinusoidal wave has the maximum amplitude of 0.6 mm, but the amplitude (A) changes as a parabolic function of radial position (r)

$$A = 0.6 - \frac{(r - 61.5)^2}{100}$$

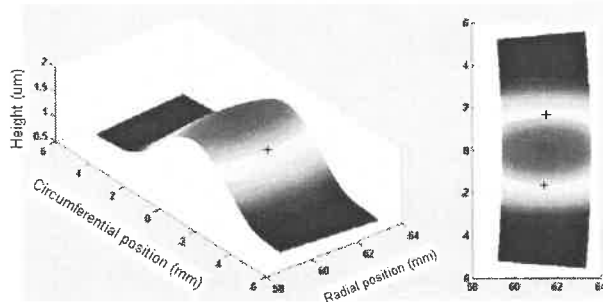


FIGURE 6. Details of the fiducial feature. The theoretical contact points with the v-groove is indicated by the crosses.

The two points of contact are shown in Figure 6 at a radial position 61.5 mm and a phase of 60° from the peak along the cosine shape. The contacts are 2.7 mm apart and 0.3 mm below the top of the feature. The slope at the contact point is 22°, but the maximum slope is 25.22°. The slope was selected to be larger than the critical angle for a pair of mating aluminum surfaces. This was measured to be 9° or a friction coefficient of 0.16. A customized diamond tool with a 27° clearance angle was supplied by Chardon Tool to avoid hitting the back of the tool on the down slope of the cosine wave. This tool was used for fabricating both the optical and fiducial surfaces.

TOOLPATH GENERATION

The generation of the toolpath for the optical and the fiducial surfaces are complicated by their NRS shape, the need for feedforward acceleration command to minimize the following error [2] and compensation for tool nose radius.

Axes Selection

A 2-axis diamond turning machine was used for fabrication with the part mounted on a spindle on the Z-axis and FLORA II moving on the X-axis (radial direction on the part). As a result, z motion can come from the Z-axis, the FLORA II

or both. Because the range of motion of FLORA is sufficient to do the entire surface, the Z-axis was not used in the fabrication.

Tool Nose Radius Correction

The shape of the optical surface and fiducial feature can be represented by a point cloud after correction for the tool nose radius. This correction, which must be calculated over the entire NRS surface, is illustrated in Figure 7 for the case of a tilted flat. As the part rotates, the normal to the surface changes and with it the contact point of the tool and workpiece. For any part radius (R_p) and θ , the contact angle can be calculated and the corrections δx and δz in Figure 4 can be found. A new set of commands (x, z) will result.

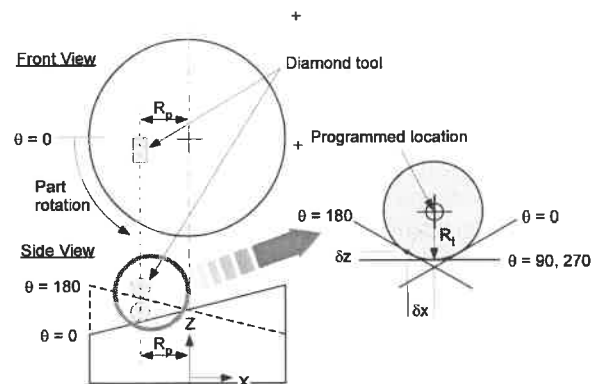


FIGURE 7. Tool Nose Radius correction

Optical Surface

After finding the new tool path with tool radius (0.9977 mm) compensation for the M4 surface in Figure 1, a uniform polar mesh grid of 50x361 (radius, mm x spiral dist., mm) is created and a bilinear interpolation algorithm [3] is applied in real-time to generate position and acceleration commands as $f(r, \theta)$ at the update rate of 20 KHz.

Figure 8 shows the tool motion at the spindle speed of 500 rpm and cross-feed rate of 200 mm/min while the radial position changes from the outer radius of the optic (49 mm) to zero. The tool positioning error in the center of travel at 90 sec is less than ± 100 nm. Figure 8 also provides a close-up view of the tool motion and error magnitude for two revolution of spindle at 90 sec. The dominant motion amplitude is at the second harmonic of spindle speed because of the biconic surface.

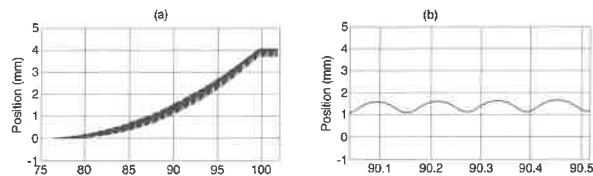


FIGURE 8. (a) Toolpath motion for entire M4 surface vs. $t(\text{sec})$ and (b) close-up view for toolpath at 90 sec for two spindle revs.

Fiducial Surface

For the fiducial feature, the position command and acceleration command can be generated from the sinusoidal function. A polar mesh grid similar to the optical surface is created including the tool radius compensation. Since the dominant frequency to create the fiducial surface is 50x that of the optic, the spindle speed was reduced to 21 rpm. This reduction in spindle speed made the dominant frequency to create the optical and the fiducial features the same.



FIGURE 8. Machined M4 mirror with fiducials

SURFACE MEASUREMENT

Figure 8 shows the finished mirror with fiducial features, Figure 9 the shape of the optical surface and Figure 10 the setup for using a CGH to measure the biconic. The diffractive features on the CGH modify the spherical wavefront from the Zygo GPI such that it will be normal to the M4 mirror and the reflected wave will match the incoming spherical wave.

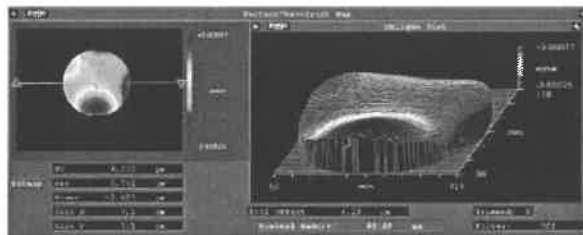


FIGURE 9. Optical surface has RMS shape error of $0.7 \mu\text{m}$.

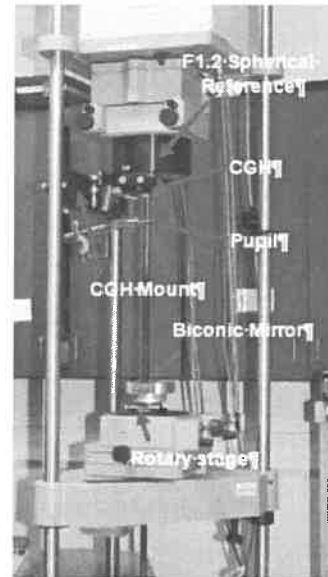


FIGURE 10. CGH setup for biconic mirror

CONCLUSIONS

The FLORA II has been integrated to a DT machine and used to produce freeform optical surfaces with accompanying fiducial surfaces for assembly. The mechanical design, high-resolution encoder and high-speed control system have demonstrated a breakthrough technology to fabricate and assemble multi-element optical systems.

ACKNOWLEDGEMENT

Principal funding is through the NSF Grant DMI-0556209 monitored by G. Hazelrigg. PMDi provided software support for the control system and Chardon tool provided the diamond tool.

REFERENCES

- [1] Zdanowicz, E., *Design of a Fast Long Range Actuator*, MS Thesis, Dept of MAE, NC State University, 2009.
- [2] Chen, Q., *Design and Control of a Fast Long Range actuator for Single Point Diamond Turning*, PhD Dissertation, Dept of MAE, NC State University, 2009.
- [3] Q. Chen, T. Dow, K. Garrard and A. Sohn, Derek Richardson, 2007. "A Fast Long Range Actuator", Proc. of ASPE, Vol 43, 2008.
- [4] Chen, Q., T. Dow, T.A., Garrard, K and Sohn, A., 2007. "A Fast Long Range Actuator", Proc. of ASPE, Vol 42, pg 287-290, 2007.
- [5] Garrard, Ken and Sohn, Alex, "Off-axis biconic mirror fabrication", PEC Annual Report, Vol 20, January 2003, pg 263-286.

HIGHER BANDWIDTH, BETTER PRECISION ! ???

Jan van Eijk^{1,2}

¹Department PME, Faculty 3ME
Delft University of Technology
Delft, The Netherlands

²Mechatronic Innovation and Concept Engineering bv
Eindhoven, The Netherlands

INTRODUCTION

In the paper **Dynamic Design for Improved Control** presented at the ASPE annual meeting in 2009 it was assumed that the performance of the motion device will improve when the bandwidth of the position control is increased. In mechanical positioning systems one often drives a mass with a force while the displacement is used for feedback control. In this case the "Zero dB"-crossover frequency in the Open-Loop transfer function can be used to indicate the "Power" of the control. A higher "bandwidth" indicates a system better capable of handling disturbances.

In the paper changing the design of the mechanical structure to optimize the dynamics was used to allow for a higher feedback gain and obtain a better performance.

This assumption has been made over decades in the development of high performance motion systems. Yet the textbooks on control design generally claim that a trade-of between performance and "control effort" must be made and that optimal control does not always means to achieve the highest bandwidth. Three examples will be given of systems where it seems to be good to reduce the bandwidth to obtain better performance. These cases are;

1. Point of Control is not measured
2. Noise in the Sensor
3. Playability in Optical Disc Drives

Using examples from the fields of Optical Disc Drives, Magnetic Bearings and Semiconductor XY-stages this issue will be discussed and the major boundary conditions that must be taken into account will be described.

Application of these findings in the case of the demonstrator used in the previous paper will be

presented. The results indicate that in this case the obtained improvements are significant.

POINT OF CONTROL IS NOT MEASURED

In general it is a good practice to use measurements as close as possible to the real point of interest in the system. A clear example that violates this principle is the case where an encoder is placed at a motor while the point of interest is driven through a mechanical construction. Traditional rotary motors with a transmission to the load are good examples. A simplified model of such system is shown in figure 1.

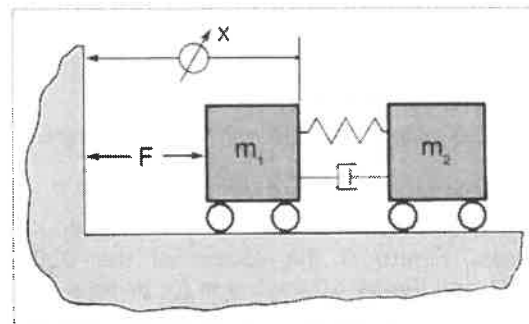


Figure 1. Schematic of a system driven by a motor with an encoder. The mass to be positioned is not directly measured.

Due to the co-location of driving force and sensor no limitations to the bandwidth come from the dynamics. Thus a high bandwidth can be obtained. In that case the position of the motor is locked with a large stiffness. As a result the internal mode of vibration will prove to be not damped by the controller and any disturbance, like set-point motion, will lead to an un-damped response. In this case practical tuning rules have been derived to assure that damping from the controller is also damping the internal vibration.

Here a reduced bandwidth clearly leads to better performance.

NOISE IN THE SENSOR

Measuring at a point away from the point of interest is one example of measurement limitations. When a sensor is used that has a significant amount of measurement noise something similar may occur. In a project on development of a magnetically suspended platform for Optical Disc Mastering an a-synchronous error of about 3 nm had to be obtained. When the platform rotates each subsequent track on the disc must be equally spaced. Error contributions that repeat every revolution are less important.

The sensors used for measuring the platform position proved to have a significant noise. Giving the controller a high bandwidth will mean that the mechanical system will follow the erroneous measurement information up to a high frequency. As the standard deviation of the error is obtained by adding the contributions over the full frequency band of interest adding noise over a larger frequency range will lead to larger errors. As a result the design used a lower bandwidth and additional integrators. The integrators helped to improve the tracking performance for low frequencies. The lower bandwidth reduced the accumulation of sensor noise to actual error motions.

PLAYABILITY IN OPTICAL DISC DRIVES

Increasing the bandwidth of the feedback controllers in Optical Disc Drives has long been limited by dynamic effects. Initially a bandwidth of about 700 Hz was needed. With the introduction of CD-ROM and later formats these requirements has gone up to about 3500 Hz and higher. It proved to be very helpful that the dimensions of lenses were reduced over time.

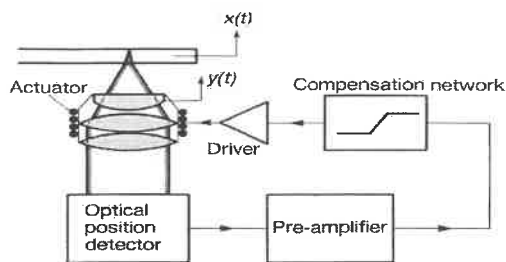


Figure 2. Schematic of the focus control for Optical Disc Player.

In the initial years there were some aspects mentioned that suggested that lower bandwidth would be desirable. The first notion is that a high bandwidth results in large control actions and thus energy consumption. This is clearly a general statement and not extremely applicable in this case. The disc actuators use the control to follow the moving track on the disc, both vertically and radially. When track disturbances are present the control forces will make the actuator move with the same accelerations as the track is moving. (Typical track movement is 200 μm and error 0.2 μm) Higher bandwidth controllers will reduce the errors from 0.2 μm to 0.1 μm but the extra force needed will be about 0.05% of the nominal forces and the additional energy is negligible.

A second issue that is mentioned relates to the "Playability" of the system. With this term the capability to reproduce slightly damaged optical discs is defined. The assumption is that a damaged disc may lead to errors in the measurement signal seen by the optical system. When the signal is lost the control will give a command to the actuator to move in some direction. The higher bandwidth system will do this more aggressively because of the higher stiffness in the controller. In reality the maximum acceleration is bounded by the maximum current in the actuator and this non-linearity makes the reaction independent of the selected bandwidth for the controller.

EXAMPLE SYSTEM RESULTS

In the paper presented last year a linear motor driven wafer stage was considered. A schematic representation is shown in figure 3. The wafer is positioned on the top of the mirror block. The sides of this block are the measurement surface for laser interferometers that measure the stage motion. The stage is vertically supported by an air-bearing and driven in horizontal direction by a linear motor.

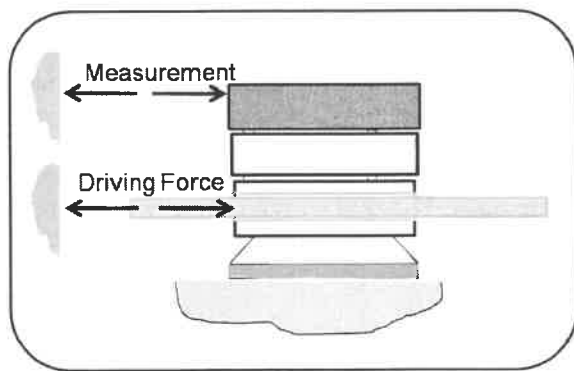


Figure 3. Schematic presentation of the wafer stage.

As the driving force is applied below the Centre of Gravity, CoG, of the module and measurement is done above the CoG resonance of the module on the air-bearing will lead to a negative phase shift around 200 Hz. This is illustrated in figure 4.

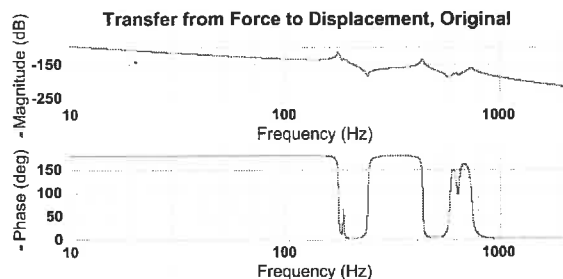


Figure 4. Open Loop Transfer function showing the resonance at about 200 Hz with its associated negative phase shift.

Due to the presence of this resonance the bandwidth for this system has to be limited to about 60 Hz. As a result the disturbance reduction at 10 Hz will be about 1/36. Higher rejection of disturbances is desired and therefore adjustments of the dynamics of the design were made. In the resulting system a resonance around 400 Hz proved to be the limiting factor. As a result the bandwidth can be raised to about 120 Hz, leading to disturbance rejection at 10 Hz of about 1/144.

In this case the measurement is done at the wafer level and the noise in the interferometer is small relative to the required precision. As a consequence the optimization of the bandwidth in this case leads to the best final performance.

CONCLUSION

Going for the maximum bandwidth in feed-back control will lead to the best precision in positioning devices when the sensor signal used is faithfully representing the position to be controlled. When mechanical structures are placed between the sensor and the Point of Interest and when the sensor signal is containing significant noise other optimization targets must be used.

In the wafer stage used for dynamic modifications optimizing bandwidth has resulted in the best obtainable precision.

A SYSTEMS APPROACH TO MODELING PIEZORESISTIVE MEMS SENSORS

Robert M. Panas, Michael A. Cullinan, and Martin L. Culpepper
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts

SYNOPSIS

Piezoresistive sensing systems have characteristics that enable them to act as fine-resolution, high-speed force and displacement sensors within MEMS. High-performance piezoresistive sensing systems are often difficult to design due to tradeoffs between performance requirements, e.g. range, power, bandwidth, and footprint. Given the complexity of the tradeoffs, traditional approaches to system design have primarily focused upon optimizing a few, rather than all, elements of the sensing system. This approach leads to designs that underperform, often by a significant margin. In this paper, we present a model that captures all significant noise sources, thereby enabling a systematic approach to the design of piezoresistive sensor systems. Improvements can be made in the performance of piezoresistive MEMS sensors using the knowledge of the tradeoffs between design parameters.

INTRODUCTION

Piezoresistors are widely used in MEMS sensor systems due to their low cost, small size, low phase lag, and large dynamic range. They have been used to create MEMS nanomanipulators [1], biocharacterization instruments [2], pressure sensors [3], inertial sensors [4], mass sensors [5], and elements of high-speed atomic force microscopes (AFMs) [5,6,7]. One must understand and manage the tradeoffs between size, bandwidth, resolution, power, and dynamic range in order to obtain high-performance in piezoresistor-based sensor systems. This requires the ability to accurately predict noise sources and how their effects propagate through a sensing system. In this paper we present a systems approach to understanding and managing dominant noise sources.

A thorough study of noise sources adds certainty to the design process, and can yield systems that have sensing which is competitive with capacitive and optical sensor systems. For MEMs applications, this means that one can use

cost- and size-appropriate sensors. For example, AFM systems with piezoresistive cantilevers are smaller, have better sensitivity and higher bandwidth than those which use photoelectric sensors [5].

The three most common types of MEMs piezoresistors are single crystal silicon, polysilicon and metal film piezoresistors. Single crystal silicon piezoresistors typically have the highest dynamic range due to their high gauge factors (20 to 100 depending on doping concentration [8, 9]) and low flicker noise. The gauge factor of single crystal silicon depends upon crystallographic orientation [10], therefore this material is typically only used in single axis, cantilever-type force sensors [6, 8, 9]. For multi-axis devices, polysilicon and metal piezoresistors are typically used given the gauge factor is fairly isotropic. [11]. Polysilicon piezoresistors tend to have a lower gauge factor (10-40 depending on doping [10]) and higher flicker noise than single crystal silicon due to the effect of grain boundaries [12, 13]. Metal film piezoresistors have a significantly lower gauge factor (~ 2) than single crystal and polysilicon piezoresistors but also have lower flicker noise due to their higher carrier concentration [5].

PIEZORESISTIVE SENSOR SYSTEM MODEL

System Architecture

We have created a model that describes the noise dependent performance of the DC piezoresistive sensor system that is depicted in Fig. 1. This model is generic in the sense that the sensor may be used to measure a force, F , that is applied to the compliant element or a displacement, δ , of the compliant element. The sensor system is composed of a voltage source, V_s , which energizes a span temperature compensated (STC) Wheatstone bridge. An instrumentation amplifier boosts the signal from the bridge, which is nulled with a bias voltage and read by an Analog-to-Digital Converter

(ADC). This circuit is commonly used and adds little noise to the strain signal.

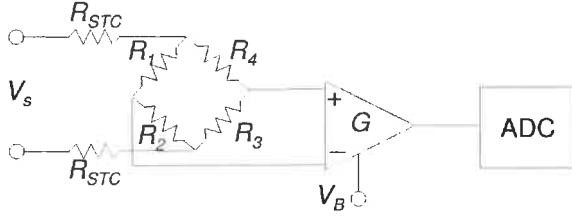


FIGURE 1: Schematic layout of DC piezoresistive sensor system.

Model Form

Thermal, electrical and mechanical noise sources are incorporated into the model as either measured or predicted values. Partial derivatives of the full model yield the sensitivity of the signal output, Ψ_M , to noise sources. The noise spectrum of the system is then obtained by considering the combined effect of all noise sources. The model may be considered in several sections, as indicated by the boxes in Fig. 2. The boxes correspond to the domains within Fig. 1. We apply the following inputs to the flexure: (1) Force or displacement signal, Ψ , (2) mechanical environmental noise, e.g. vibrations, σ_{Mv} , and (3) thermomechanical displacement noise, σ_{Mt} , which has a spectral density, $S_{Mt}(f)$.

$$S_{Mt}(f) = 4k_B T \left(\frac{2\zeta}{k\omega_n} \right) \quad (1)$$

In Eq. (1), k_B is Boltzmann's constant, T is temperature in Kelvin, k is the stiffness, ζ is the damping ratio, f is the noise frequency in Hz, and ω_n is the natural frequency of the compliant structure in radians per second [14]. Ambient mechanical and thermal noise spectral densities were measured in our laboratory.

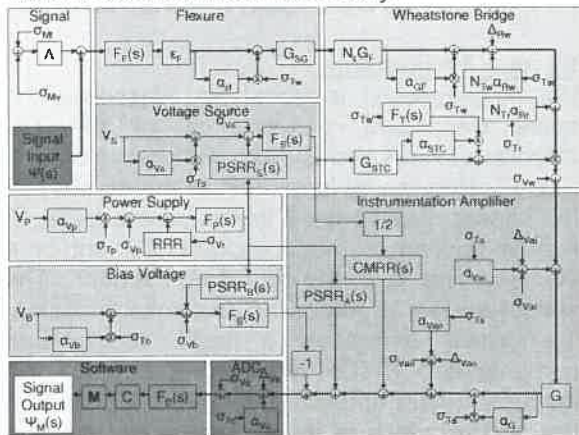


FIGURE 2: Model representation with electrical and thermal noise sources.

The mechanical displacement noise sources are scaled by a constant, Λ , to match the units of the signal, where Λ has either a unity value for displacement signals or value of k for force signals. The compliant structure, i.e. flexure, acts as a non-dimensional mechanical filter, $F_F(s)$, to the signal as expressed by the Laplace transform of the dynamics of the compliant structure. All filters, $F(s)$, in the model are assumed to have unity steady state gain. The compliant structure also acts as a mechanical transducer that converts force and displacement into strain. This is represented in the model as the flexure gain, ϵ_F . The expressions for rectangular flexures with common types of boundary conditions are provided in Table 1. L_f , b_f , and h_f are the length, width and height of the flexure. The variable, E represents Young's Modulus and N_b is the number of flexures that mechanically act in parallel to form the compliant structure.

TABLE 1: Common forms of the flexure gain, ϵ_F .

	Fixed-Guided	Fixed-Free
δ	$3h_f/L_f^2$	$3h_f/(4L_f^2)$
F	$3L_f/(N_b b_f h_f^2 E)$	$6L_f/(N_b b_f h_f^2 E)$

The flexure transfer function exhibits a thermal sensitivity $\alpha_{\epsilon F}$, when E is sensitive to changes in Wheatstone bridge temperature, σ_{Tw} . The gain, G_{SG} , accounts for the averaging of the strain field over the finite geometry of the piezoresistor in the length and thickness dimensions.

$$G_{SG} = \left(1 - \frac{L_r + 2L_o}{\gamma L_f} \right) \left(1 - \frac{h_r}{h_f} \right) \quad (2)$$

In Eq. (2), L_r is the length and h_r is the height of the piezoresistor. The sensor is placed a distance L_o from the ground of the flexure. The constant γ , has value of 1 for fixed-guided or 2 for fixed-free flexures. The second term is normally neglected for surface mounted or deposited piezoresistors since h_r is usually much smaller than h_f .

The signal is transformed from the mechanical to electrical domains via the Wheatstone bridge. The bridge sensitivity to strain depends upon the bridge strain type, N_ϵ , which is defined as the number of strain sensitive resistors in the bridge. The gauge factor, G_F , exhibits thermal sensitivity α_{GF} . The thermal sensitivity is defined by the number of thermally active resistors on the bridge, N_{Tw} , and the temperature coefficient of

resistance, α_{Rw} . Likewise for other resistors forming the bridge but located elsewhere, N_{Tr} and α_{Rr} . The imbalance in the bridge resistors is represented by the dimensionless term Δ_{Rw} . The output signal is scaled by the bridge energizing voltage V_S . The sensor noise, σ_{Vw} , is composed of Johnson and flicker noise contribution from each of the 4 piezoresistors. This noise exhibits spectral density $S_{Vw}(f)$.

$$S_{Vw}(f) = 4k_B T R + \frac{1}{16} \sum_{i=1}^4 \frac{V_S^2 \alpha_i}{C_C \Omega_i f} \quad (3)$$

In Eq. (3), R is resistance, α is the Hooge constant, C_C is the carrier concentration, and Ω is the volume of the piezoresistor [7, 15]. The bridge voltage is attenuated by the gain of the STC, G_{STC} , which has thermal sensitivity, α_{STC} .

$$G_{STC}(T) = \frac{\frac{G_{STC}}{R_W}}{\frac{R_W}{R_W + R_{STC}}} \left[1 + \frac{\frac{\alpha_{STC}}{R_{STC}}}{\frac{R_{STC}}{R_{STC} + R_W} (\alpha_{Rw} - \alpha_{Rstc}) T} \right] \quad (4)$$

The STC attenuates the bridge energizing voltage through the use of resistors, R_{STC} , with high thermal sensitivity, α_{Rstc} , in series with the bridge resistance R_W . These resistors are used to cancel the thermal variation in the signal due to α_{GF} and α_{eF} [16].

$$R_{STC} = \frac{R_W (\alpha_{GF} + \alpha_{FI})}{\alpha_{Rstc} - \alpha_{Rw} - \alpha_{GF} - \alpha_{FI}} \quad (5)$$

The STC and bridge resistors are separated by some distance; therefore they may experience different temperatures. The transfer function $F_T(s)$ captures this by modeling the system as a thermal low pass filter. The thermal variations occurs at relatively low frequencies, therefore the bandwidth of $F_T(s)$ is large enough to approximate as unity.

The bridge energizing voltage has a thermal sensitivity, α_{GF} , to changes in its temperature, σ_{Ts} . Electrical noise from the energizing voltage, σ_{Vs} , and power supply noise as attenuated by the power supply rejection ratio $PSRR_S(s)$ are added to the signal. A filter, $F_S(s)$, is used to reduce the noise from the voltage source. The bias voltage is treated as equal to the bridge energizing voltage.

The instrumentation amplifier gain, G , has a thermal sensitivity α_G to the amplifier temperature, σ_{Ta} . The gain is calculated by constraining the maximum input to the ADC to a fraction of its operating range, u , which is generally 0.9, or 90% of the ADC's voltage range, V_{range} .

The maximum signal is found by substituting in the flexure's yield stress, σ_y , as modified by a safety factor, η , that protects against failure of the flexure's material. After the preceding we have:

$$G = \frac{u V_{range} E \eta}{2 \sigma_y G_{SG} N_e G_F G_{STC} V_S} \quad (6)$$

The signal is biased by V_B at the output of the amplifier. The error terms associated to noise (σ_{Vah} , σ_{Vao}), offsets (Δ_{Vah} , Δ_{Vao}) and the thermal sensitivities of the offsets (α_{Vah} , α_{Vao}) are added to the signal. The variations in the common mode of the bridge output are attenuated by the common mode rejection ratio, $CMRR(s)$. The same is done for power supply variations through the power supply rejection ratio, $PSRR_A(s)$.

The variation in the supply voltage, V_P , is caused by thermal sensitivity, α_{Vp} , to the supply temperature, σ_{Tp} , electrical noise σ_{Vp} and ripple at the input to the supply circuit, σ_{Vr} . The ripple is attenuated by the ripple rejection ratio RRR. A filter, $F_P(s)$, should be used to reduce the noise from the power supply.

The output voltage signal is read in from the amplifier to an ADC. The ADC adds electrical noise, σ_{Vc} , thermal sensitivity, α_{Vc} , to the ADC temperature, σ_{Tc} , and an offset Δ_{Vc} to the signal.

The digital signal is passed through a digital filter, $F_D(s)$, which may be adjusted to attenuate noise outside of the signal spectrum. The signal is scaled by a calibration coefficient, C , which is found by enforcing equality between Ψ_M and Ψ ,

$$C = \frac{1}{\epsilon_F G_{SG} N_e G_F G_{STC} V_S G} \quad (7)$$

When multiple sensors are used to obtain multi-axis measurements, a coordinate transform matrix, $\mathbf{A}$, is used to transform the vector of sensor readings, $\mathbf{p}$, into a vector of the values in the axes of interest, $\mathbf{v} = \mathbf{A}\mathbf{p}$. Uncorrelated noise from each sensor is attenuated by the averaging effect of combining the multiple sensor readings, captured by $\mathbf{M}$, which may be written as a vector to calculate the performance of the j th axes of interest.

$$\mathbf{M}_j = \sqrt{\sum_k \mathbf{A}_{j,k}^2} \quad (8)$$

Dominant Noise Sources and System Characteristics

Partial derivatives for the dominant noise sources, σ_{V_w} , $\sigma_{V_{ai}}$, σ_{T_w} , are listed below.

$$\begin{aligned}\frac{\partial \Psi_M(s)}{\partial \sigma_{V_w}} &= M C F_D(s) G \\ \frac{\partial \Psi_M(s)}{\partial \sigma_{V_{ai}}} &= M C F_D(s) G \quad \frac{\partial \Psi_M(s)}{\partial \sigma_{V_{ao}}} = M C F_D(s) \quad (9) \\ \frac{\partial \Psi_M(s)}{\partial \sigma_{T_w}} &= M C F_D(s) G G_{STC} F_S(0) V_S [\Delta_{Rw} F_T(s) \alpha_{STC} \dots \\ &+ N_T \alpha_{Rw} + \Psi_F(0) G_{SG} N_i G_F (\alpha_{GF} + \alpha_{ei} + F_T(s) \alpha_{STC})]\end{aligned}$$

The system is evaluated using spectral analysis where all noise sources are considered as unbiased, uncorrelated, and normally distributed with spectral densities, $S_n(f)$. The spectral densities are scaled by their respective frequency dependent sensitivities and geometrically summed to obtain the system noise spectral density, $S_{\Psi_m}(f)$.

$$S_{\Psi_m}(f) = \sum_n \frac{\partial \Psi_M(2\pi if)}{\partial \sigma_n} \left(\frac{\partial \Psi_M(2\pi if)}{\partial \sigma_n} \right)^* S_n(f) \quad (10)$$

INSIGHTS FROM THE MODEL

Electronic Sources

The model is generalized so that it may be used with a wide range of applications. Through this model, we may gain insight on best design of general and specific sensor systems. The model assumes the use of high-performance electrical components – instrumentation amplifier (Analog Devices AD624), voltage source and bias (Texas Instruments REF50xx series), and ADC (National Instruments 9215 ADC). This is essentially a best practice that ensures that these electronics are not a significant source of noise. There relevant noise values for these components are provided in the references. We will shortly show that sensor noise is the dominant noise source in well-designed sensing systems; therefore AC bridges are typically not needed to reduce amplifier noise.

When the resistance of the piezoresistor is low, noise from the instrumentation amplifier may become significant. The input noise from the amplifier is added to the signal before the signal is amplified, therefore reducing the signal output by the bridge may make the amplifier input noise the dominant source. If the signal is small, thermomechanical noise, i.e. random vibration in the beam caused by energy exchange between the flexure and environment, may become a

major noise source once propagated via the electronics.

Mechanical Sources

The mechanical noise sources do not significantly contribute to the overall noise in most sensor systems because this is attenuated by physical filters (e.g. via optical tables) before they reach the measurement signal.

Thermal Sources

Errors caused by thermal fluctuations can generally be avoided by proper system design. The Wheatstone bridge may be thermally balanced by placing the bridge resistors close together so that they are subject to the same temperature. Similarly, STC resistors may be used to make the gauge factor and flexure transfer function effectively thermally insensitive. Bridge offsets generated by manufacturing inaccuracies are compensated for by the bias voltage. The thermal sensitivity of the bridge offset, however, is unaffected by the bias voltage as may be discerned from Eq. (9). In cases where the manufacturing inaccuracies are large, it may become necessary to minimize thermal fluctuations in the environment through the use of insulation or active temperature controls. In most cases, this type of thermal control is not necessary since relative manufacturing inaccuracies are typically small in MEMS. In most cases, noise in the piezoresistor itself limits the resolution of the sensor system.

Johnson and Flicker Noise

The noise in the sensor may be broken into two sources: (1) Johnson noise caused by the thermal agitation of electrons in a conductor and (2) Flicker noise caused by conductance fluctuations that manifest during the capture and release of charge carriers in the piezoresistor [15]. Doping concentration affects resistivity, gauge factor and carrier concentration of silicon piezoresistors, therefore silicon piezoresistors may be Johnson or flicker noise dominated. There is a tradeoff between noise and sensitivity as dopant concentration is varied.

In the case where the performance of the sensor is limited by flicker noise, an optimal sensor length and thickness will exist. As the length and thickness of the sensor increases, the sensor volume and therefore number of carriers increases. This acts to decrease flicker noise. The average amount of strain in the sensor also

decreases as the length and thickness of the sensor increase. In balancing these two effects, optimal length and thickness may be found. The optimal sensor length is $y/3$. The fixed-guided condition and other boundary conditions are often found in multi-axis flexures. The optimal sensor to flexure thickness ratio is $1/3$, which is consistent with prior force sensor work [7].

PIEZORESISTIVE SENSOR DESIGN

Reduced Piezoresistive Sensor System

Model

One of the most important parameters describing sensor performance is the dynamic range, i.e. the ratio of range to resolution of the system. The range and resolution are functions of the flexure geometry but the dynamic range is typically dependent only on the piezoresistor itself. Therefore, it is generally best to optimize the sensor system to achieve the highest practical dynamic range. From the model it was determined that the three largest noise sources were the Johnson noise, flicker noise and instrumentation amplifier noise. In the simplified model, only these three noise sources are passed through the system model to create a reduced expression for the resolution of the sensor. The dynamic range, DR, of the sensor is given in Eq. (11),

$$DR = \frac{\sigma_y N_e G_F G_{STC} V_S \left(1 - \frac{L_r + 2L_o}{\gamma L_r}\right) \left(1 - \frac{h_r}{h_f}\right)}{\eta EM \sqrt{4k_b TRB + \frac{1}{16} \sum_{i=1}^4 \frac{\alpha_i V_S^2}{C_{Ci} \Omega_i} \ln(r) + S_{vai} B}}, \quad (11)$$

$$R = \frac{\rho N_r^2 L_r}{b_r h_r} \quad \Omega = L_r b_r h_r$$

where ρ is the resistivity, B is the bandwidth, S_{vai} is the spectral density of input noise of the instrumentation amplifier, and N_r is the number of turns in the resistor. The bandwidth of the noise may be written as a function of the signal frequency where the pole of the software first order, low pass filter is placed at a multiple r of the signal frequency, f_{sig} . The approximation of this bandwidth is given by Eq. (12) [15].

$$B = \frac{\pi}{2} (r - 1) f_{sig} \quad (12)$$

This simplified model makes it possible to optimize the dynamic range of the sensing system for most cases. However, when very small forces or displacements are being measured, the thermomechanical noise may become greater than the noise from the instrumentation amplifier.

EARLY RESULTS AND DISCUSSION

The noise characteristics of a simple quarter bridge ($N_e=0.25$) polysilicon piezoresistive sensor was compared to model predictions as shown in Fig. 5. The sensor and electronics are shielded from external noise sources. The sensor is located on a large aluminum thermal reservoir within a Faraday cage. The flicker noise characteristics of the polysilicon piezoresistive sensor were experimentally determined. The spectral density of the noise was measured from 0.01 Hz to 5 kHz, corresponding roughly to the common range of operation for such sensors.

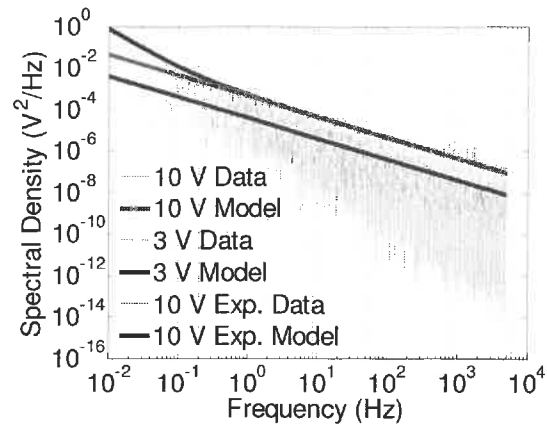


FIGURE 5: Polysilicon Piezoresistive sensor noise spectrum compared to predictions.

The model indicates that the sensor flicker noise should be the dominant source over the full range of measurement when the bridge is energized at 10 V. This prediction is verified by the measured spectral density. The predicted and measured noises are 77 mV and 78 mV respectively. The model also correctly predicts the change in noise spectral density resulting from a reduction in the bridge energizing voltage from 10 V to 3 V. In the reduced voltage scenario, the predicted and measured noises are 23 mV and 21 mV respectively.

In the third scenario studied in the experiment, the electrical and thermal shielding surrounding the polysilicon piezoresistor was removed to deliberately exposed the sensor to random temperature variations ('Exp. Data'). The spectral density of these temperature variations was measured and propagated through the system model to predict the effect of exposing the sensor on the noise spectral density. The electrical noise prediction was unaffected by this change, however the thermal noise component of the prediction rose significantly to become a

dominant source over the low frequencies (0.01 to 1 Hz) as shown in Fig 6. This effect was observed in the measurements of the spectral densities with and without thermal shielding. This indicates that thermal effects on system noise can effectively be integrated into a cohesive model as described in the previous sections.

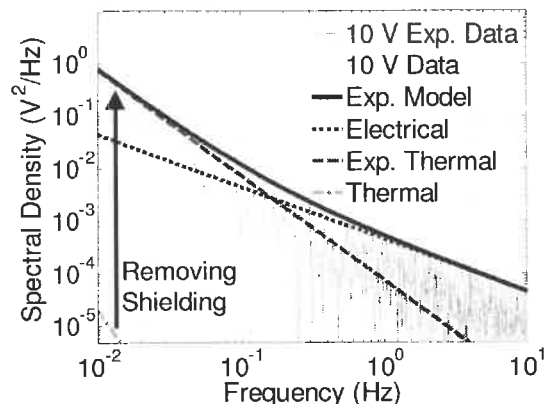


FIGURE 6: Measurement of noise spectral densities with and without thermal shielding.

REFERENCES

- [1] Yang Y, Dong Z, Qu Y, Li M, and Li W, A Programmable AFM-Based Nanomanipulation Method Using Vibration-Mode Operation. Proc. 4th Int. Conf. Nano/Micro Engineered and Molecular Systems, 2008, 681-685.
- [2] Otero J and Puig-Vidal M, Low-noise Instrumentation for the Measurement of Piezoresistive AFM Cantilever Deflection in Robotic Nanobiocharacterization Applications. IEEE Int. Instrumentation and Measurement Technology Conf. 2008, 1392-1396.
- [3] Kanda Y and Yasukawa A. Optimum Design for Silicon Piezoresistive Pressure Sensors. Sensors and Actuators A, 1997; 62: 539-542.
- [4] Yazdi N, Ayazi F, and Najfi K. Micromachined Inertial Sensors. Proc. IEEE. 1998; 86: 1640-1659.
- [5] Li M, Tang H and Roukes M. Ultra-sensitive NEMS-based Cantilevers for Sensing, Scanned Probe and Very High-Frequency Applications. Nature Nanotechnology, 2007; 2: 114-120.
- [6] Thysen J, Boisen A, Hansen O, and Bouwstra S. Atomic Force Microscopy Probe with Piezoresistive Read-out and a Highly Symmetrical Wheatstone Bridge Arrangement. Sensors and Actuators A, 2000; 3: 47-53.
- [7] Harley J and Kenny T, 1/F Noise Considerations for the Design and Process Optimization of Piezoresistive Cantilevers. Journal of Microelectromechanical Systems, 2000; 9: 226-235.
- [8] Yu X, Thaysen J, Hansen O, and Boisen A. Optimization of Sensitivity and Noise in Piezoresistive Cantilevers. Journal of Applied Physics. 2002; 92: 6296-6301.
- [9] Pruitt B and Kenny T. Piezoresistive Cantilevers and Measurement System for Characterizing Low Force Electrical Contacts. Sensors and Actuators A. 2003; 104: 68-77.
- [10] French P. Polysilicon: A Versatile Material for Microsystems. Sensors and Actuators A. 2002; 99: 3-12.
- [11] Cullinan M, Panas R, and Culpepper M. Design of Micro-scale Multi-Axis Force Sensors for Precision Applications. Proc. Annual Meeting of the American Society for Precision Engineering. 2009.
- [12] Upreti N and Singh S. Grain Boundary Effect on the Electrical Properties of Boron-doped Polysilicon Films. Bull. Mater. Sci. 1991; 14: 1331-1341.
- [13] Berderlow R, Webber W, Dahl C, Schmitt-Landsiedel D, and Thewes R, Low-Frequency Noise of Integrated Poly-Silicon Resistors. IEEE Trans. Electron Devices. 2001; 48: 1180-1187.
- [14] Ramos D, Tamayo J, Mertens J, Calleja M, Villanueva L, and Zaballos A. Detection of bacteria based on the thermomechanical noise of a nanomechanical resonator: origin of the response and detection limits. Nanotechnology, 2008; 19: 035503.
- [15] Senturia S, Microsystem Design. Kluwer Academic Publishers, Boston: 2002.
- [16] Omega Engineering, Inc., The Pressure Strain and Force Handbook. Stamford, CT: 2006.

OPTIMAL MULTIPLE LINEAR MOTOR COMMUTATION BY REAL-TIME OPTIMIZATION

Shalom D. Ruben and Tsu-Chin Tsao
Mechanical and Aerospace Department
University of California, Los Angeles
Los Angeles, California

INTRODUCTION

Precision motion control is becoming increasingly more important in the areas of nanotechnology. Lithography, for one, is requiring greater precision as transistor density continues to increase. Permanent magnet synchronous linear motors (PM-SLM) are among the most prevalent in precision machines due to their long-range travel, high force density, and high precision and accuracy. Lack of mechanical coupling needed, versus rotary motors, and in conjunction with an air bearing eliminate many disturbances such as backlash and friction [1].

We extend optimal minimum power commutation for a single n -phase Halbach linear motor to multiple linear motor systems, such as the Sub-Atomic Measuring Machine (SAMM) [2], having four 6-phase halbach linear motors and Multi-Alignment and Positioning System (MAPS) [3], having four 3-phase halbach motors. A Halbach linear motor produces an independent levitation force along with the thrust force [4].

A new optimal commutation algorithm is proposed where minimum power is desired subject to a heat symmetric solution. Minimizing the temperature gradient, will also minimize deformation across the system and would be beneficial in nano-manufacturing systems such as MAPS. The heat symmetric solution is not closed-form and a second-order cone programming (SOCP) problem needs to be solved at each servo cycle. We develop our own SOCP solver that will iterate to a solution within a phase cycle of $110\mu s$ and implement on Delta Tau's Power PMAC real-time control system.

MINIMUM POWER

The force produced by the linear motor is proportional to the current supplied to the n coils and is defined as

$$f = Bi \quad (1)$$

With the motor law defined as the mapping from current to force ($B : i \rightarrow f$), the inverse motor law (or commutation) which maps a desired force to current ($T_{inv} : f \rightarrow i$) is derived. As seen by the control loop in Fig. 1, the controller C generates a desired force and must be transformed, by T_{inv} , into a desired current to be sent by the motor amplifier.

The minimum power commutation problem is a

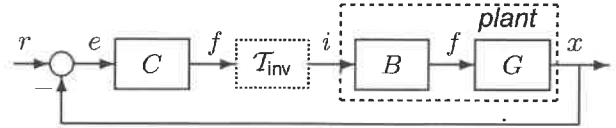


FIGURE 1. feedback control loop for a linear motor

least-norm problem defined as

$$\begin{aligned} &\text{minimize} && i^2 \\ &\text{subject to} && f = Bi \end{aligned}$$

The solution is known to be $i^* = B^\dagger f = B^T(BB^T)^{-1}f$. In ideal motors, as in (2) where all the phases are perfectly sinusoidal, it is known as sinusoidal commutation and is simplified to

$$i^* = \frac{2}{na^2} B^T f = \frac{2}{na^2} (f_x C_x + f_z C_z)$$

where $BB^T = \frac{na^2}{2} I$.

We now describe the minimum power commutation for a multiple linear motor system such as MAPS in Fig. 2. The local force vector f for a multiple motor system with m motors will be defined as

$$\begin{aligned} f &= [f_{x1} \ f_{z1} \ f_{x2} \ f_{z2} \ \dots \ f_{xm} \ f_{zm}]^T \\ &= [f_1 \ f_2 \ \dots \ f_m]^T \end{aligned}$$

and the relationship between the local motor forces and the global forces ($F = [F_x \ F_y \ T_z \ F_z \ T_x \ T_y]^T$) is

$$F = Af$$

$$\begin{bmatrix} f_x \\ f_z \end{bmatrix} = a \begin{bmatrix} \cos(\omega x - \phi) & \cos(\omega x - \phi + \frac{\pi}{n}) & \cdots & \cos(\omega x - \phi + \frac{\pi(n-1)}{n}) \\ \cos(\omega x - \phi + \frac{\pi}{2}) & \cos(\omega x - \phi + \frac{\pi}{n} + \frac{\pi}{2}) & \cdots & \cos(\omega x - \phi + \frac{\pi(n-1)}{n} + \frac{\pi}{2}) \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_n \end{bmatrix} \quad (2)$$

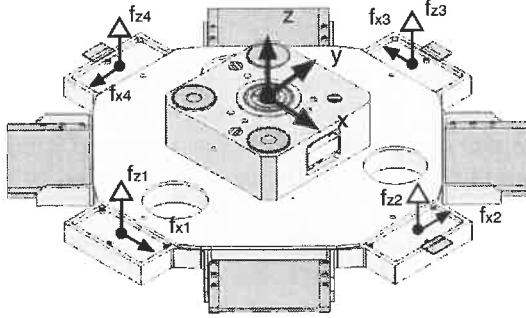


FIGURE 2. 4 linear motors produce 8 forces to control 6 degrees of freedom

From Fig. 2, the relationship between F and f can be shown to be

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ R & 0 & R & 0 & R & 0 & R & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & -R & 0 & 0 & 0 & R & 0 & 0 \\ 0 & 0 & 0 & -R & 0 & 0 & 0 & R \end{bmatrix} \quad (3)$$

Where R is the distance from the center of the platen to the center of each motor. The matrix $A \in \mathbf{R}^{6 \times 8}$ has two vectors in its nullspace. This means that there are two non-zero local force vector f that produce zero global force F . If $f_{x1} = -f_{x2} = f_{x3} = -f_{x4}$ or $f_{z1} = -f_{z2} = f_{z3} = -f_{z4}$ then there is a zero net-force ($F_x = F_y = T_z = F_z = T_x = T_y = 0$) acting on the platen.

The relationship between the current vector i and the local forces for a multiple linear motor system is

$$B = \begin{bmatrix} B_1(x_1) & & & 0 \\ & B_2(x_2) & & \\ & & \ddots & \\ 0 & & & B_m(x_m) \end{bmatrix} \quad (4)$$

where

$$B_k(x_k) = \begin{bmatrix} C_x(x_k)^T \\ C_z(x_k)^T \end{bmatrix}$$

and x_k is the local position of the stator and mover of each motor. For MAPS, the local positions of each motor are

$$\begin{aligned} x_1 &= x + R \sin(\theta_z); & x_2 &= y + R \sin(\theta_z) \\ x_3 &= -x + R \sin(\theta_z); & x_4 &= -y + R \sin(\theta_z) \end{aligned}$$

The motor law for a 6-DOF multiple linear motor system is defined as

$$F = ABi \quad (5)$$

Where the current vector i is defined as

$$\begin{aligned} i &= [i_{11} \ i_{12} \ \dots \ i_{1n} \ i_{m1} \ i_{m2} \ \dots \ i_{mn}]^T \\ &= [i_1 \ \dots \ i_m]^T \end{aligned}$$

The minimum power commutation is found by solving

$$\begin{aligned} &\text{minimize} \quad i^2 \\ &\text{subject to} \quad F = ABi \end{aligned} \quad (6)$$

The solution to this problem is given as $i^* = (AB)^\dagger F$. It is possible to attack this problem in two stages, a kinematic stage where the desired local force vector is found $f = A^\dagger F$ and an electro-magnetic stage, where the minimum power current vector, given the kinematic stage, is found $i = B^\dagger f$ giving us $i = B^\dagger A^\dagger F$.

In general, this two-stage method does not give the same solution since $(AB)^\dagger \neq B^\dagger A^\dagger$ [5]. In fact, the power solved by the one stage method is less than or equal, and therefore optimal, to the solution of the two stage method ($(AB)^\dagger F \leq B^\dagger A^\dagger F$). Therefore, separating the kinematic stage and the electro-magnetic stage when minimizing the power is not in general minimal.

Due to the structure of the ideal linear motors (namely that $BB^T = (na^2/2)I$), the one stage solution turns out to be equal to the two stage solution ($(AB)^\dagger F = B^\dagger A^\dagger F$). It is important to note that using the two stage method, and separating the kinematic and electro-magnetic stages, in real systems won't give the minimum power solution because it is not possible to manufacture each motor and each coil to be exactly the same. See Fig. 3 for a complete block diagram of MAPS.

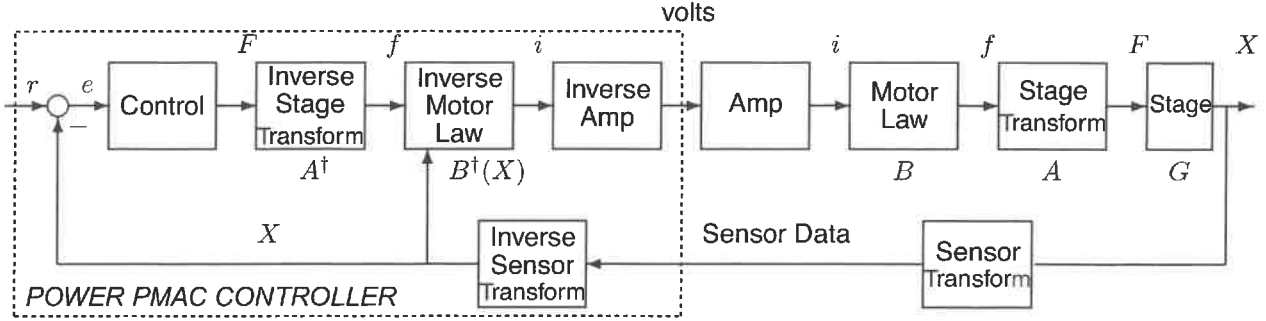


FIGURE 3. MAPS CONTROL LOOP SHOWING TRANSFORMATIONS

MINIMUM TEMPERATURE GRADIENT

We now propose a new commutation problem, by adding the constraints that each motor's power (i_k^2), separately, is equal to one another:

$$\begin{aligned} &\text{minimize} && i^2 \\ &\text{subject to} && F = ABi \\ &&& i_k^2 = p; \end{aligned} \quad (7)$$

This problem is motivated by the temperature gradient, and subsequent non-symmetric deformation of the platen, produced by the motors generating different power profiles. By each motor generating the same power we force heat symmetry on the platen, and therefore symmetric deformation about the center of the platen. This problem is non-convex due to the added quadratic equality constraints [6].

The above stated problem is difficult to solve. We propose to solve this problem by a 2-stage method. The first step of this method involves solving a convex problem that minimizes the maximum power that any of the four motors can have ($\min \max i_k^2$). The power for the motor with the greatest power is defined as P . The second step is to project the solution of the min max problem onto the set of solutions that each motor has power equal to P . This projection is possible by adding a non-zero current in the nullspace of B (see table 1). The projection onto the space of equal power motors is defined by

$$\text{Proj}(i_o, P) = \begin{bmatrix} i_{o1} + \alpha_1 i_{N1} \\ i_{o2} + \alpha_2 i_{N2} \\ \vdots \\ i_{on} + \alpha_n i_{Nn} \end{bmatrix} \quad (8)$$

TABLE 1. Convex and Algebraic solution to the non-convex problem (7)

Given: desired global force vector F

Stage 1.

$$\text{solve} \quad P = \text{argmin}(\max_k(i_k^2))$$

$$\text{subject to} \quad F = ABi$$

Stage 2.

$$\text{Project } i: \quad i^* = \text{Proj}(i, P)$$

Output: current vector i^*

where α_k is defined as

$$\alpha_k = -\frac{i_{ok}^T i_{Nk} \pm \sqrt{i_{ok}^T i_{Nk} - i_{Nk}^T i_{Nk} (i_{ok}^T i_{ok} - P)}}{i_{Nk}^T i_{Nk}} \quad (9)$$

Remember that $i_{Nk} \in \mathcal{N}(B_k)$, which means that any added current in this direction can increase the power without changing the global desired force ($F = AB(i_o + \alpha i_N) = ABi_o + \alpha ABi_N = ABi_o$). By setting the desired power P equal to the power given by new current vector ($(i_o + \alpha i_N)^T (i_o + \alpha i_N)_k = P$), α_k can be found by solving the following quadratic equation.

$$\alpha_k^2 (i_{Nk}^T i_{Nk}) + 2\alpha (i_{ok}^T i_{Nk}) + i_{ok}^T i_{ok} - P = 0 \quad (10)$$

The proof that the method in table 1 solves the non-convex problem is as follows:

Proof. Let the solution of (7) give P^* for each motor and the solution for the min-max problem of table 1 be P_{hs} . By the problem in (7), the total power satisfies $4P^* \leq 4P_{hs}$. On the other hand, by the problem in table 1, $P_{hs} \leq P^*$. Therefore, $P^* = P_{hs}$. $\square$

TABLE 2. Solving for i where $F_x = 1; T_z = 4R$

Method	i_1	i_2	i_3	i_4	i
$\min P$	.6	.27	.06	.27	1.2
Heat Sym	.4	.4	.4	.4	1.6

The results of a simulated example, where the desired global forces are $F_x = 1, T_z = 1$ and all the rest are zero, are shown in Table 2. Two solutions, a minimum power (6) and a heat symmetry (7), are presented. Both solutions produce the same force, while the power output per motor is quite different.

REAL-TIME OPTIMIZATION SOLVER

Many convex optimization problems involve the solution of a sequence of problems until an optimal solution has been reached. Unless your update rate is slow, as in the chemical process industry [7], then embedded convex optimization methods are not viable. Control of mechanical systems with update rates in microseconds, such as MAPS with an update rate of $110\mu s$, have traditionally not been able to use embedded convex optimization techniques. With the speed of computers increasing, these problems are within reach even for update rates in microseconds. The following section describes the development of a solver using interior-point method, specifically for our minimum temperature gradient problem described in the previous section. By writing our own code, we will be able to take advantage of every exploitable feature that a general solver might not be able to find.

First, a reduction of our problem size is possible. Assuming that $BB^T = (na^2/2)I$ holds, the min-max problem in table 1 becomes equivalent to solving the following problem:

$$\begin{aligned} &\text{minimize} && \max_k (-f_k) \\ &\text{subject to} && F = Af \end{aligned} \quad (11)$$

and the desired current i is found by solving the least-norm problem $i = B^\dagger f^*$. In this section, we will only deal with solving the convex optimization problem of (11), giving f^* , but keep in mind that to get the total solution $i^* = \text{Proj}(B^\dagger f^* | P = \max_k (B^\dagger f^*)_k)$ must be calculated. In words, the solution (f^*) from (11) is found, then the current (i) that minimizes the power subject to $f^* = Bi$ is solved. Finally the motor currents are projected so that their power is equal.

The problem of (11) can be reformulated as an

SOCP (12).

$$\begin{aligned} &\text{minimize} && q \\ &\text{subject to} && F = Af \\ &&& -f_k \leq q \end{aligned} \quad (12)$$

Both equality and inequality constraints need to be dealt with in (12). It is possible to eliminate the equality constraint from, by defining f as

$$f = f_o + Zy$$

Where $f_o = [f_{ox_1} \ f_{ox_2} \ \dots \ f_{ox_m} \ f_{oz_m}]^T$ is a feasible solution ($Af_o = F$) and $Z = [Z_1 \ Z_2 \ \dots \ Z_m]^T$ spans the nullspace of A . By eliminating the equality constraint from (12) we get

$$\begin{aligned} &\text{minimize} && q \\ &\text{subject to} && -(f_o + Zy)_k \leq q \end{aligned} \quad (13)$$

and reduce the number states in (12), $-q \ f \in \mathbb{R}^8$, from nine to three, $-q \ y = [y_1 \ y_2]^T$. This leaves an inequality constrained problem. Primal-Dual methods, advanced Interior point methods [6], are used to solve inequality constrained problems and are prevalent in state-of-the-art software such as SeDuMi, and SDPT3. A simpler interior-point method, known as the primal log-barrier method, is used to solve our SOCP problem. Although primal-dual methods are more advanced than the primal barrier method, the details shown here can easily be extended to the primal-dual method. The idea behind the log-barrier method is to remove the inequality constraints leaving you with an unconstrained optimization problem (14).

$$\text{minimize} \quad tq + \sum_k \phi_k(q, y) \quad (14)$$

Where $\phi_k(x) = -\log(q^2 - (f_o + Zy)_k^T (f_o + Zy)_k)$, is convex and differentiable everywhere, can be thought of as a weight that increases exponentially the closer you get to having an infeasible solution (a barrier). The barrier method is an iterative method that solves (14) many times for different values of t . Each solution, starting with a small value of t and updating ($t = \mu t$) at each iteration, is known as the central path. As t increases the affect from the log-barrier function decreases so the solution approaches the solution of the original primal problem (13).

The barrier method has two loops, the outer loop described by the central path and the inner which solves the newton system of (14). Solving the

newton system, a set of linear equations, is the most time intensive part of the algorithm. Defining our state vector $\mathbf{u} = [\mathbf{y} \ \mathbf{q}]^T$ we rewrite (14) as

$$\text{minimize } t\mathbf{c}^T\mathbf{u} + \sum_k \phi_k(\mathbf{u}) \quad (15)$$

and define the newton system as

$$H\Delta\mathbf{u} = -\mathbf{g} \quad (16)$$

where

$$\mathbf{g} = t\mathbf{c} + \nabla_{\mathbf{u}}\phi \quad (17)$$

$$H = \nabla_{\mathbf{u}}^2\phi \quad (18)$$

Each newton iteration, the newton direction $\Delta\mathbf{u} = -H^{-1}\mathbf{g}$ is solved and used to update the state vector by

$$\mathbf{u}^{(j+1)} = \mathbf{u}^{(j)} + s\Delta\mathbf{u} \quad (19)$$

where s is chosen such that the new cost is less than the previous cost,

$$t\mathbf{c}^T\mathbf{u}^{(j+1)} + \sum_k \phi_k(\mathbf{u}^{(j+1)}) < t\mathbf{c}^T\mathbf{u}^{(j)} + \sum_k \phi_k(\mathbf{u}^{(j)}) \quad (20)$$

Iterations of the inner loop continue until the stopping condition of $\lambda^2 \leq \epsilon$ ($\lambda^2 = \mathbf{g}^T H^{-1} \mathbf{g}$) is met or it is forcibly stopped by defining the max number of inner iterations.

Depending how s is chosen can increase complexity in each inner iteration. Backtracking line search is an iterative process where s is initialized at one and updated ($s = \beta s$) until (20) is met. A very simple, closed-form, way to calculate s , but not as aggressive as backtracking line-search, is

$$s = \frac{1}{1 + \lambda} \quad (21)$$

What you gain in simplicity and speed of calculation, you lose in the number newton iterations it takes to converge. Therefore, backtracking line-search will be more costly to perform but will reduce the number of newton iterations and therefore possibly the total time to solve the problem. In the results section, we will compare these two ways of calculating s .

To find the inverse of the hessian, H , can be costly unless you can find a closed-form equation for the inverse. Since H is only a 3×3 matrix the inverse can be found as a closed-form set of equations. It could also be useful, to enhance stability, to calculate the closed-form equation for the Cholesky factorization (L) of H and calculate

TABLE 3. Solving for i where $F_x = 1$; $T_z = 4R$

Force Stop	Line Search	Chol-esky	Number Iterations	Time per Solve(μ s)
	✓		32	30.375
	✓	✓	32	34.453
			37	22.578
		✓	37	28.266
✓	✓		20	21.797
✓	✓	✓	20	25.156
✓			20	12.406
✓		✓	20	15.454

the inverse by forward and backward substitution to $H^{-1} = L^{-T}L^{-1}$. The matrix L is also useful for the stopping criterion λ^2 and backtracking line search since λ^2 is also needed. Where

$$\lambda^2 = \mathbf{g}^T H^{-1} \mathbf{g} = \mathbf{g}^T L^{-T} L^{-1} \mathbf{g} = L^{-1} \mathbf{g}^2 \quad (22)$$

RESULTS AND CONCLUSIONS

Table 3 compares optimality gap, number newton iterations and time to solve of the following adaptations of the SOCP solver: To calculate optimality gap, we first solved a specific problem ($F_x = 1$; $F_y = 2$; $T_z = 3$; $F_z = 4$; $T_x = 5$; $T_y = 6$) using CVX and used this solution to compare optimality gap with our SOCP solver running on an 800MHz laptop PC. To get a good estimate, the time it took to solve this problem 10^6 times is recorded.

The terms μ , β , and α , used in line-search and outer update were all kept constant. The results show that it is possible to use convex optimization in real-time applications. Since MAPS controller updates the phase every 110μ s, we can use any of the adaptations of the SOCP solver. The fastest solves were forced to stop before converging and not using backtracking line-search or the Cholesky factorization. This makes sense since all of these take time to calculate. You'll also notice that using the line-search algorithm will reduce the number of newton iterations, as assumed.

The algorithm was implemented on Delta Tau's Power PMAC controller. The Power PMAC has separate memory locations for user-written control and phase algorithms in C++ and by default servo is updated at 2.5KHz while phase is 9KHz. Fig. 4 shows the power of each motor while MAPS was under closed-loop regulation and minimum power commutation. An external disturbance was added at motor #1(see Fig. 2), making motor #3 a pivot point where power is the

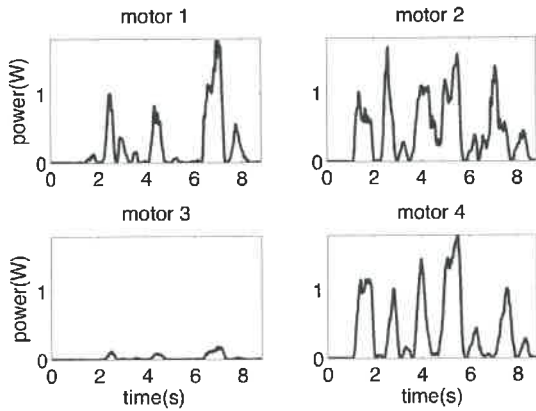


FIGURE 4. minimum power implementation

smallest of all the motors. The implementation of the minimum temperature gradient commutation is shown in Fig. 5, where disturbance was also added at motor #1. Notice from Fig. 6 that implementing either commutation scheme has no effect on the following error. In this paper we devel-

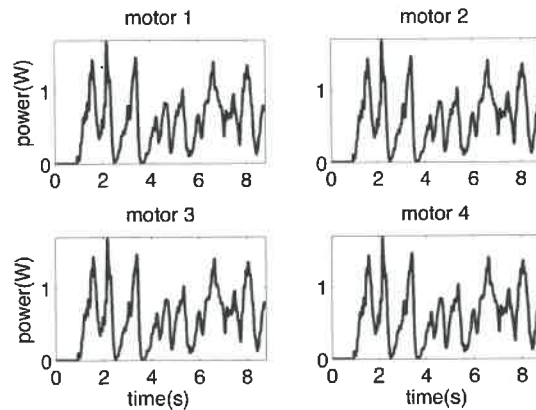


FIGURE 5. minimum temperature gradient implementation

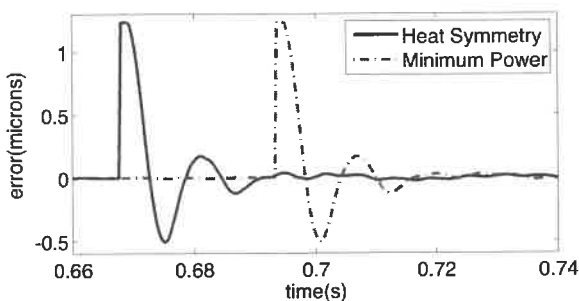


FIGURE 6. following error to a step command
oped two optimal commutation schemes for multiple linear-motor precision systems. The first was a minimum power commutation which resulted in

a closed-form solution, that is easily implemented in real-time. The second was a minimum temperature gradient commutation which resulted in needing to solve SOCP problem. We wrote our own SOCP solver that was able to solve the problem in less than 35 microseconds, which is within one phasing cycle of $110\mu s$. Finally implementation results of minimum power and minimum temperature gradient commutation were presented.

ACKNOWLEDGMENTS

The authors wish to acknowledge Prof. L. Vandenberghe (UCLA, Electrical Engineering Department) for his invaluable advise. The authors would also like to thank Delta Tau Data Systems, Inc., for donating the Power PMAC real-time control hardware and software. This work was supported in part by the National Science Foundation under grant no. DMI 0327077 and CMMI 0751621.

REFERENCES

- [1] Kiong TK, Heng LT, Sunan H. Precision Motion Control: Design and Implementation. Springer; 2008.
- [2] Holmes M, Hocken RJ, Trumper D. The long-range scanning stage: a novel platform for scanned-probe microscopy. Precision Engineering. 2000;24:191–209.
- [3] Ruben SD, Tsao TC, Hocken RJ, Feserman RR, Ozturk O, Brien J, et al. Mechatronics and Control of a Precision Motion Stage for Nano-Manufacturing. In: Proceedings of ASME Dynamic Systems and Control Conference; 2009. .
- [4] Trumper DL, Kim WJ, Williams ME. Design and Analysis Framework for Linear Permanent Magnet Machines. IEEE Transactions on Industry Applications. 1996;32(2):371–379.
- [5] Laub AJ. Matrix Analysis for Scientists and Engineers. SIAM; 2004.
- [6] Boyd S, Vandenberghe L. Convex Optimization. Cambridge University Press; 2004.
- [7] Mattingley J, Boyd S. 1. In: Automatic Code Generation for Real-Time Convex Optimization. Cambridge; 2009. Chapter of Convex Optimization in Signal Processing and Communications.

FUNDAMENTALS OF FRICTION MODELING

Farid Al-Bender¹

¹Department of Mechanical Engineering, Division PMA
Katholieke Universiteit Leuven
Leuven, Belgium

INTRODUCTION

Friction modeling has been steadily gaining in interest over the last few decades. However, owing to the complexity of the friction phenomenon, no comprehensive, practicable friction model that shows all of the experimentally observed aspects of friction force dynamics in one formulation is available. Most available friction models are, in essence, empirical, that is, based on limited observations and interpretations. In this sense, the resulting models are valid only for the specific scope of test conditions, such as the level and type of excitation, used to obtain the data. On the other hand, development of simulation models and, where possible, predictive theories, at scales from atomic, through continuum, to useful engineering models, can fill empty gaps in the toolboxes available to designers and analysts.

Besides the field of tribology, where the origin of friction is one of the main topics, modeling and compensation of friction dynamics are treated in several other domains. In the machining and assembly industry, demand for high-accuracy positioning systems and tracking systems is increasing. Research on controlled mechanical systems with friction is motivated by the increasing demand for these systems. Friction can severely deteriorate control system performance in the form of higher tracking errors, larger settling times, hunting, and stick-slip phenomena. In short, friction is one of the main players in a wide variety of mechanical systems.

This communication presents an overview of friction model-building, which starts from the generic mechanisms behind friction to construct models that simulate observed macroscopic friction behavior. First, basic friction properties are presented. Then, the generic friction model is outlined. Hereafter, simple heuristic/empirical models are discussed, which are suitable for quick simulation and control purposes. An example of these is the Generalized Maxwell-Slip model.

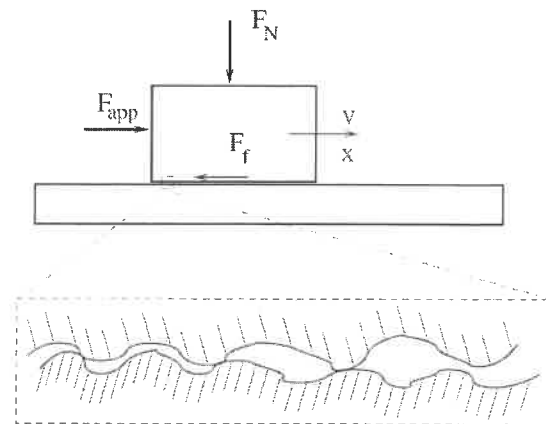


FIGURE 1. Basic friction configuration.

BASIC FRICTION BEHAVIOR

Considering friction as a mechanical system, (see Fig.1), a close examination of the sliding process reveals two friction regimes, namely, the pre-sliding regime and the gross sliding regime, (see Fig.2). In the pre-sliding regime the adhesive forces, owing to asperity contacts, are dominant, and thus the friction force is primarily a function of displacement rather than velocity. The reason for this behavior is that the asperity junctions deform elasto-plastically, thus behaving as nonlinear hysteretic springs.

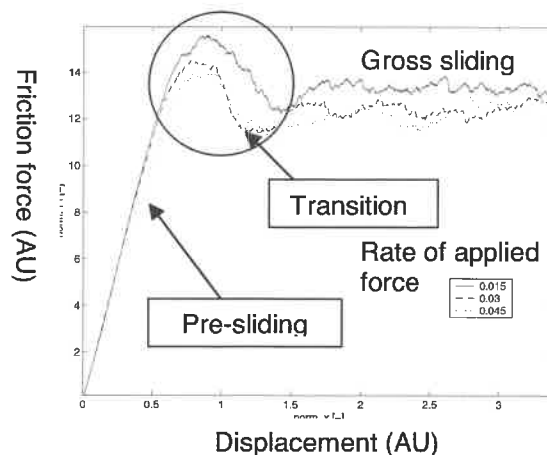


FIGURE 2. The two friction regimes and the transition between them.

As the displacement increases, more and more junctions break and have less time to reform, resulting eventually in gross sliding.

The sliding regime is, thus, characterized by a continuous process of asperity junction formation and breaking such that the friction force becomes predominantly a function of the velocity [1]. The transition from pre-sliding to gross sliding is a *criticality* that depends on many factors such as the relative velocity (to be envisaged as the displacement rate) and acceleration of the sliding objects, see [2].

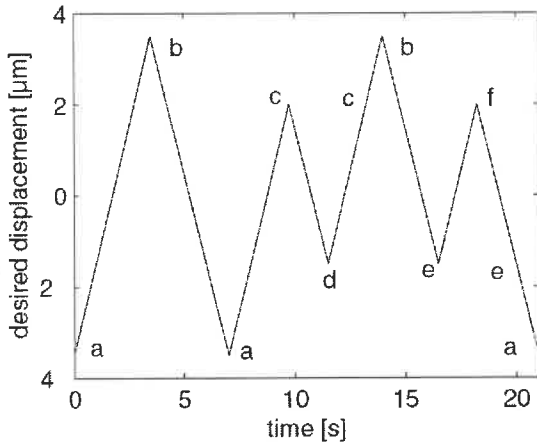


FIGURE 3. Example of desired motion in the pre-sliding regime.

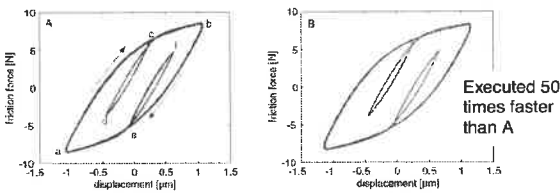


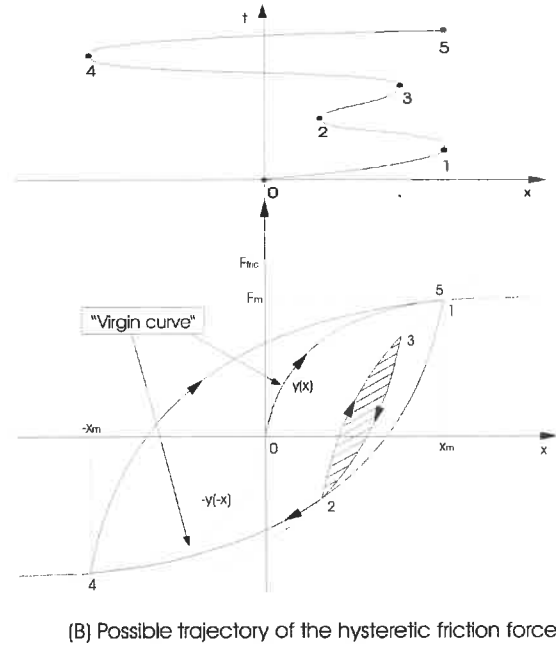
FIGURE 4. Hysteresis behavior as a result of the trajectory of Fig. 3.

Pre-sliding behavior

At very small displacements, that is, in the pre-sliding regime, experiments reveal a hysteretic displacement-dependent friction force [3,4]. When a pre-sliding displacement command, such as that shown in Fig. 3, is applied to the block, the force-displacement behavior of Fig. 4 results. The position signal is chosen such that there is an inner loop within the outer hysteresis loop. The resulting friction-position curve is rate independent (compare the right and left panels of Fig. 4). In other words, the friction-position curve is independent of the speed of the applied position signal. When an inner loop is closed, (c-

d-c), the curve of the outer loop (a-c-b) is followed again, proving the nonlocal memory characteristic of the hysteresis. The shape of the hysteresis function is determined by the distribution of the asperity heights, the tangential stiffness, and the normal stiffness of the contact.

This hysteresis behavior arises primarily from micro-slip, that is, the breaking of adhesive contacts, just as in the Maxwell-Slip model discussed further below. The contribution of deformation losses, which are hysteresis losses in the bulk materials, depends on the relative value of this part as compared to the adhesive part, as well as on the tangential stiffness of the asperities, which governs the extent of deformation before slip.



(B) Possible trajectory of the hysteretic friction force

FIGURE 5. Characteristics of hysteresis with nonlocal memory.

Figure 5 explains the constitution of hysteresis according to Masing's rules. The force-displacement curve initially follows:

$$F_{\text{virgin}} = f(x)$$

$$\text{with } f(x) = \begin{cases} y(x) & ; x \geq 0 \\ -y(-x) & ; x \leq 0 \end{cases}$$

Upon reversal at any point x_m , a doubly dilated version of the virgin curve is followed:

$$F_{\text{fric}} = F_m + 2f\left(\frac{x - x_m}{2}\right)$$

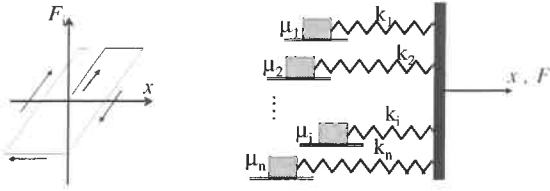


FIGURE 6. Modeling of hysteresis using Maxwell-Slip elements.

This behavior can also be modeled discretely by a parallel connection of Maxwell-Slip elements [5], as shown in Fig. 6.

Gross sliding

When the asperity junctions are continually being created and broken, the frictional interface is in the gross sliding regime. Two main characteristics are of interest here. The first is the steady-state friction force behavior with increasing steady-state sliding velocities, generally known as the Stribeck curve, Fig. 7. The second is the change of the friction force with the velocity variation, known as the *friction lag* or *friction memory* phenomenon, Fig. 8.

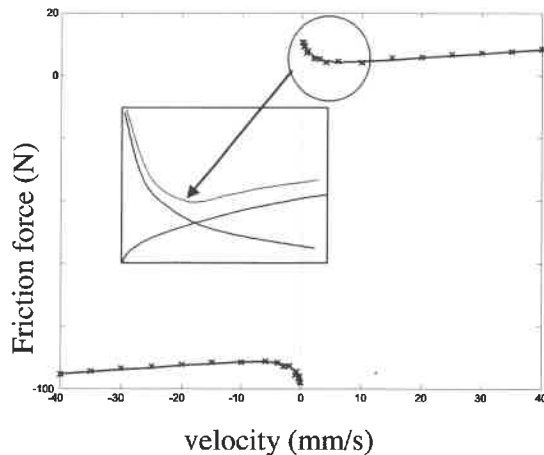


FIGURE 7. The Stribeck curve consists of velocity weakening and velocity strengthening.

The Stribeck curve

When the friction force is measured at constant velocity values (Fig. 7), the resulting functional relationship has a characteristic form. For increasing velocities, the friction force initially decreases to a minimum (*velocity weakening*) and then increases again (*velocity strengthening*). In lubricated sliding, this characteristic is known as the Stribeck curve, where the velocity weakening arises from the

initial buildup of hydrodynamic pressure, while the velocity strengthening is attributed to the viscous shear of the lubricating film.

The same behavior seems to hold true for dry friction, which justifies using the same name, that is, the Stribeck curve, to describe it. The actual form of the friction-velocity curve is determined by various process parameters, namely, the normal creep or, equivalently, the time evolution of adhesion, the surface topography, and the asperity parameters, primarily the tangential stiffness and inertia [2].

Friction lag

Friction lag, also called hysteresis in the velocity, or frictional memory, is manifested by a lag in the friction force relative to the sliding velocity. The origin of friction lag in lubricated friction relates to the time required to modify the lubricant film thickness, which is known as the *squeeze effect*. Friction lag is also observed in dry friction experiments (Fig. 8), where lubrication is not used. The mechanism causing friction lag in dry sliding is similar to that for lubricated friction, namely, that the local adhesion coefficient increases with the time of contact of two opposing asperities, owing to normal creep. In other words, time is required before the friction force changes with changing sliding velocity. Since the normal creep is caused by the sinking of the surfaces into each other, this mechanism is akin to the squeeze effect in lubricated friction. Thus, the friction force is higher for acceleration than for deceleration, so that the dynamic friction force curve circles around the steady-state curve. The Stribeck curve $s(v)$ acts as the attractor.

$$\text{Model: } \frac{dF}{dt} = f(1 - F/s(v)), \quad f(0) = 0.$$

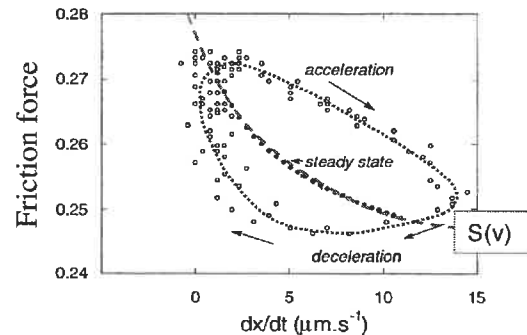


FIGURE 8. Friction lag and its constitutive equation. $s(v)$ is the velocity weakening curve also called the "Stribeck effect".

A GENERIC FRICTION MODEL

In order to reconstruct the friction behavior outlined above, in the framework of a mechanical theory, a generic model was developed [6].

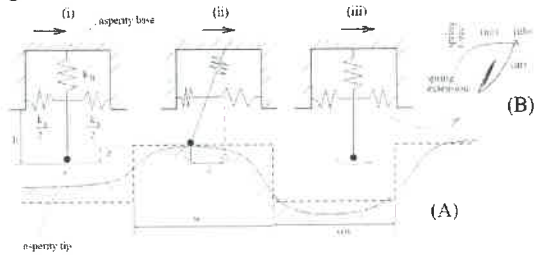


FIGURE 9. A generic representation of the sliding contact of rough bodies.

The model comprises an upper body containing point-mass asperities supported on hysteresis springs, which slides against a rigid, profiled lower surface, subject to normal creep, adhesion and deformation. The life cycle of an average asperity contact is depicted in Fig. 9. (A) An asperity is initially moving freely (i) until it touches the lower rigid surface (ii). After sticking and slipping, it breaks completely loose from the lower profile (iii). (B) depicts the hysteretic force-deformation diagram during a contact cycle of the asperity, where upon breaking loose, the asperity is assumed to dissipate all of its elastic energy through internal hysteresis losses.

Typical results of the generic model

The generic model yields the self-explanatory results reviewed in Fig. 10 through 12. The parameters used in the model have been chosen ad hoc to illustrate typical behavior, although they can also be identified from experimental results.

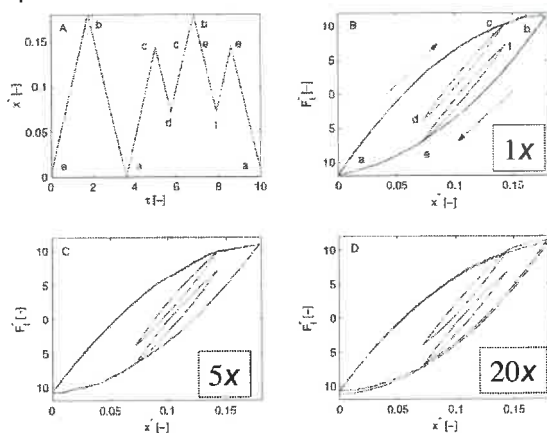


FIGURE 10. Simulation results of pre-sliding hysteresis using the generic model. Cf. Fig. 3.

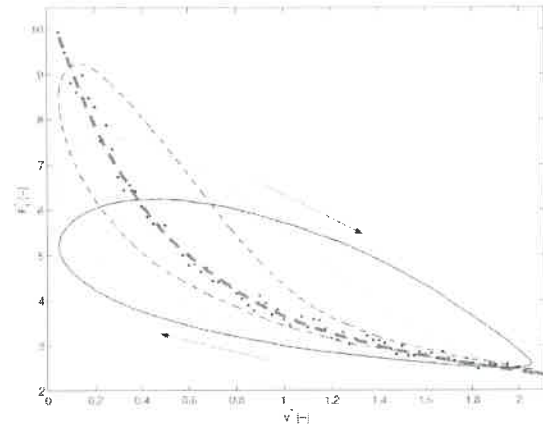


FIGURE 11. Friction lag using the generic model. The higher the acceleration, the more the loops depart from the Stribeck curve.

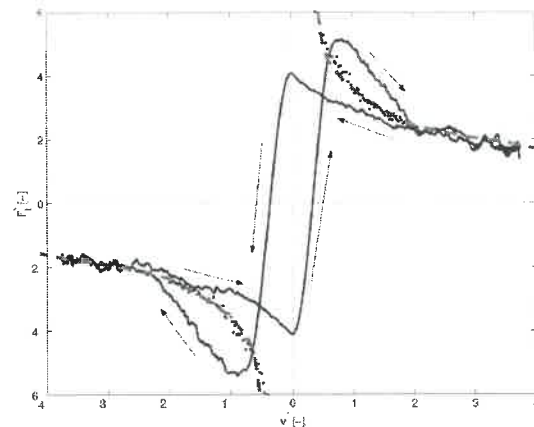


FIGURE 12. Periodic friction-velocity curves using the generic model.

Finally, simulation experiments with this model have contributed to the derivation of Generalized Maxwell-Slip (GMS) model belonging to the category of empirical models subject of the next section.

EMPIRICAL MODELS

Generalized Empirical Friction Model

Structure

Analysis of this class of model reveals that most of the existing empirical friction models correspond to a generalized friction model structure, which consists in a friction force equation and a state equation.

The friction force F_f is a generalized function of an internal state vector $\mathbf{z}$ (often representing

asperity deflections), the velocity v , and the position x of the moving object, that is,

$$F_f = f(z, v, x)$$

The state equation that describes the dynamics of the internal state vector $\mathbf{z}$ is a first-order differential equation of the form

$$\frac{dz}{dt} = g(z, v, x)$$

The functions f and g are generally nonlinear functions. In particular, it is shown that

$$f(z, v, x) = f_1(z, v, x) + f_2(v)$$

where f_1 is responsible for the transient response in the velocity, while f_2 represents the instantaneous response to velocity changes. However, this formulation corresponds to sliding and thus does not allow for pre-sliding or for passing through zero velocity.

Empirical friction modeling consists then in finding suitable expressions for the generalized functions f and g , such that the resulting model faithfully simulates all observed types of friction behavior.

Two generic conditions apply, which provide limiting conditions on the functions f and g . First, for constant velocities, the steady state ($dz/dt = 0$) friction force is a function of the velocity v only. Thus, if $v = \text{constant}$, then

$$g(z, v, x) = 0, \text{ and}$$

$$f(z, v, x) = F(v) = s(v) + f_2(v)$$

where $s(v)$ is the velocity-weakening curve, while $f_2(v)$ is the "viscous" or velocity-strengthening curve.

The second condition on the general functions is determined by the frictional behavior in the pre-sliding regime at small displacements. The friction force is then a hysteresis function of the position, with nonlocal memory characteristics [3] as given by

$$F_f = f(z, v, x) = F_h(x)$$

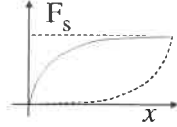
From the last two conditions, we can see that g is generally not continuous.

Evolution of the Empirical Friction Models

A possible way to evolutionarily classify existing empirical/heuristic friction models is according to the following main line.

Dahl Model

The friction force is a hysteresis function (without memory) of the displacement.

$$\frac{dF_f}{dx} = \sigma_0 \left(1 - \frac{F_f}{F_s} \text{sgn}(v) \right)$$


LuGre Model

Transforms the Dahl equation into the state equation:

$$\frac{dz}{dt} = v \left(1 - \frac{z}{s(v)} \text{sgn}(v) \right)$$

The friction force is given by:

$$F_f = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v$$

Leuven Model

This adds a hysteresis function with nonlocal memory F_h to the formulation

$$\frac{dz}{dt} = v \left(1 - \frac{F_h(z)}{s(v)} \text{sgn}(v) \right), \text{ state}$$

$$F_f = F_h(z) + \sigma_1 \dot{z} + \sigma_2 v, \text{ Friction force.}$$

Generalized Maxwell-Slip Model

This was conceived as a match between the generic, LuGre and Leuven models [7]. It is constructed by imposing a rate-state behavior to the friction blocks of the Maxwell-Slip model. Referring to Fig. 13, the state equations are:

$$\text{If stick : } \frac{dz_i}{dt} = v, \quad z_i \leq s_i(v)$$

$$\text{If slip : } \frac{dz_i}{dt} = \text{sgn}(v) C_i \left(1 - \frac{z_i}{s_i(v)} \right)$$

The friction force is given by:

$$F_f = \sum_{i=1}^{i=N} (k_i z_i + \sigma_i \dot{z}_i) + f(v)$$

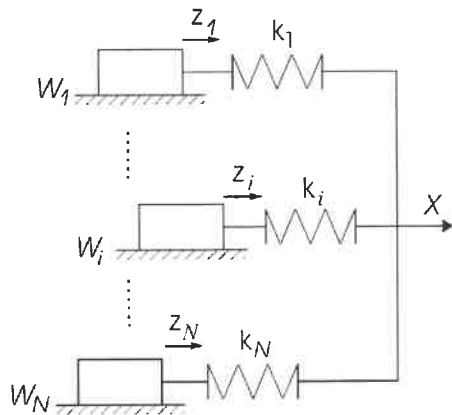


FIGURE 13. Maxwell-Slip to Generalized Maxwell-Slip model.

COMPARISON WITH EXPERIMENT

In how far the generic and GMS models agree with experimental observations depends on parameters chosen for the models. In other words, those models are effective in regard to structure. This is supported by the results of Fig. 14, which is at the same time an illustration of the stick-slip phenomenon and its intricate behavior.

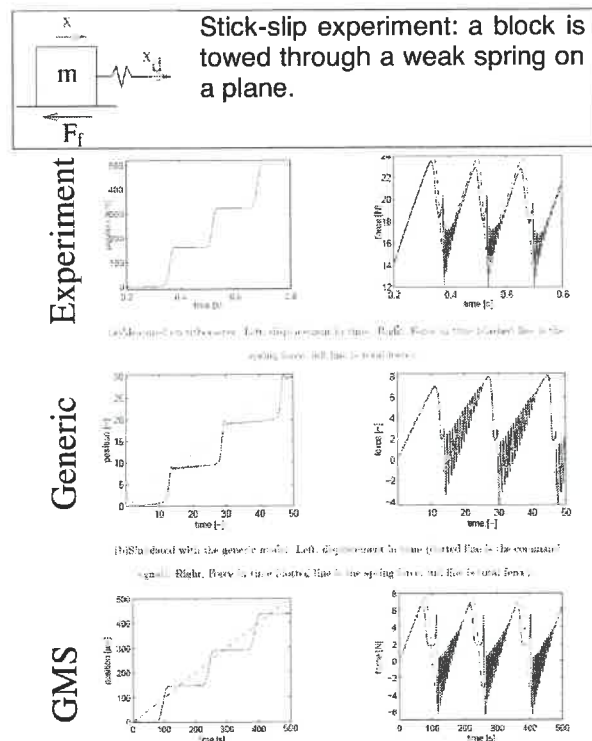


FIGURE 14. Stick-slip motion: comparison between models and experiment.

CONCLUSIONS

Friction is a complex, nonlinear phenomenon. Modeling of friction is important in many fields of science and engineering. We have reviewed a generic physical model and a number of empirical models; closing the circle by deriving the Generalized Maxwell-Slip (GMS) model, which is suitable for quick simulation and control purposes. The generic and GMS models agree well with experimental observations

Acknowledgements

This research is partially supported by the Fund for Scientific Research - Flanders (F.W.O.) under Grant FWO4283. The scientific responsibility is assumed by its authors.

REFERENCES

- [1] ArmstrongHélouvry B. Control of Machines with Friction. Norwell, MA: Kluwer, 1991.
- [2] Al-Bender F, Lampaert V, Swevers J. Modelling of dry sliding friction dynamics: from heuristic models to physically motivated models and back. CHAOS, An Interdisciplinary Journal of Nonlinear Science. 2004, 14(2): 446-460.
- [3] Swevers J., Al-Bender F, Ganseman C, Prajogo T. An integrated friction model structure with improved presliding behaviour for accurate friction compensation. IEEE Transactions on Automatic Control. 2000, 45(4): 675-686.
- [4] Lampaert V, Al-Bender F, Swevers J. Experimental characterisation of dry friction at low velocities on a developed tribometer setup for macroscopic. Tribology Letters, 2004, 16(1): 95-105.
- [5] Al-Bender F, Symens W, Swevers J, Van Brussel H. Theoretical analysis of the dynamic behavior of hysteresis elements in mechanical systems. International Journal of Non-linear Mechanics. 2004, 39: 1721-1735.
- [6] Al-Bender F, Lampaert V, Swevers J. A novel generic model at asperity level for dry friction force dynamics. Tribology Letters. 2004, 16(1): 81-93.
- [7] Al-Bender F, Lampaert V, Swevers J. The Generalized Maxwell-slip model: a novel model for friction simulation and compensation. IEEE Transactions on Automatic Control. 2005, 50(11): 1883-1887.

PRACTICAL CONTROLLER FOR PRECISION POSITIONING SYSTEM – A NEW DESIGN APPROACH AND APPLICATION TO MECHANISM WITH FRICTION

Shin-Horng CHONG¹, and Kaiji SATO¹

**¹Interdisciplinary Graduate School of Science and Engineering
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan**

INTRODUCTIONS

A variety of precision positioning system, such as machine tools, measuring machines and semiconductor manufacturing systems, generally require high performance motion controllers. For this account, a practical controller which provides simple and transparency design and structure; produces fast response, nearly zero overshoot, high robust performance and excellent accuracy is required, especially in industry.

Industry has favored classical controllers such as PID, due to their structure simplicity and well-known characteristics [1]. However, the classical controllers have met limitations when a higher positioning performance becomes more stringent and high robust characteristics are required. Many advanced controllers have been devoted to solve these problems, such as robust controllers with disturbance observer [2], time-optimal controllers [3], adaptive robust controllers [4], sliding mode controllers [5], and advanced intelligent controllers. Even though those advanced controllers are capable to gain performances satisfied in robustness of precision positioning, determination of known model parameters of the plant and sufficient knowledge of control theory in their design procedure are necessary. The performance of those controllers are based on how accurate the known model and known parameters are used in their design procedures. Their design processes are time-consuming and much labor is needed to identify its parameters. It is a difficult task for operators who are unfamiliar with advanced controllers. In industry application, these advanced controllers are often impracticable.

In this research, as a practical solution, nominal characteristics trajectory following (NCTF) controllers have been proposed and studied to overcome the occurred problems [6-13]. The controller comprises a nominal characteristic

trajectory (NCT) and a PI compensator. The NCTF controllers are able to satisfy the above mentioned system requirements to positioning where it is easy to design, adjust and do not rely on known model parameters identification. So far, the NCTF controllers have been improved steadily to enhance its performance in positioning and continuous motion control. Besides, the NCTF controllers have demonstrated better positioning performance than the Ziegler-Nichols tuning method [7]. The experimental results in [9] indicate that the Continuous Motion NCTF controller is more suitable in positioning and tracking motion controls. However, the current NCTF controllers have addressed the less accuracy of following characteristics on nominal characteristic trajectory. This phenomenon often indicates unwanted overshoot and low motion accuracy.

This paper provides a practical solution to improve the Continuous Motion NCTF controller for reduction of the overshoot and the motion error. The improved practical controller is referred to as the Acceleration Reference – Continuous Motion NCTF controller (AR-CM NCTF controller). The AR-CM NCTF controller serves in improving the following characteristic of the object motion to NCT near origin and providing the high overshoot reduction characteristics and low sensitivity to disturbance force. The AR-CM NCTF controller provides high disturbance reduction characteristics in order to achieve the improvement in positioning accuracy and the reduction of tracking error. The AR-CM NCTF controller includes the Continuous Motion NCTF controller structure. Besides the velocity reference, the acceleration reference for object motion is used as the additional controller elements. The constructed NCT and designed PI compensator of the Continuous Motion NCTF controller [10] are adopted in implementing the AR-CM NCTF controller.

IMPROVEMENT OF THE CONTINUOUS MOTION NCTF CONTROL SYSTEM

Structure

Continuous Motion (CM) NCTF control system

Figure 1 shows the structure of the Continuous Motion NCTF control system designed for high motion control performance. The controller comprises a nominal characteristic trajectory (NCT) and a PI compensator. The NCT represents the reference motion of the system and is expressed on the phase plane. The PI compensator makes the system motion follow the NCT and leads the motion to the origin of the phase plane. The NCT is constructed from the velocity and displacement responses in open-loop condition because the relationship between them is essentially influenced by the mechanism characteristics. The inclination of NCT near origin is referred as β . The NCT is used for determining the virtual error rate reference for each deviation and the PI compensator works for velocity control. The control law, u of the CM NCTF control system is

$$U(s) = \left(K_p + \frac{K_i}{s}\right)U_p(s), \quad \text{where}$$

$$U_p(s) = (X_r(s) - X(s))s - E_{NCT}(s)s.$$

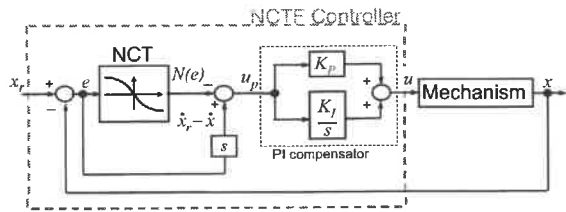


FIGURE 1. Continuous Motion NCTF control system.

Acceleration Reference – Continuous Motion NCTF (AR-CM NCTF) control system

So far, the CM NCTF controller has addressed less motion accuracy and slight overshoot problem. It is due to the object motion which does not follow the NCT accurately and this phenomenon causes overshoot near the origin on phase plane (Fig. 2). Hence, the AR-CM NCTF controller is proposed to provide the overshoot reduction characteristics and improve the motion accuracy of the system.

Figure 3 shows the AR-CM NCTF controller structure. In contrast to the Continuous Motion NCTF controller, this controller includes

additional controller elements: the δ -block and λ -block. So that, the relationship between $\dot{e}$ and $\ddot{e}$ for the object motion is considered. The control law, u of the AR-CM NCTF controller is

$$U(s) = \left(K_p + \frac{K_i}{s}\right)U_p(s), \quad \text{where}$$

$$U_p(s) = (E(s)s - E_{NCT}(s)s) + (E(s)s^2 - E_\delta(s)s^2)\lambda$$

$$\text{with } E(s)s = (X_r(s) - X(s))\frac{s}{1 + T_d s} \quad \text{and}$$

$$E(s)s^2 = (X_r(s) - X(s))\left(\frac{s}{1 + T_d s}\right)^2.$$

The δ -block is constructed from the open-loop acceleration with the inclinations near origin, δ which is same as NCT, β . It is a virtual acceleration reference for the object motion. By considering the acceleration trajectory in the control system, the object motion can be controlled and follow more accurately the NCT and ended at the origin without any overshoot on the phase plane. In improving the motion accuracy, the λ gain of the λ -block is adjusted adaptively to work well with δ -block. Overall, the additional controller elements in the AR-CM NCTF controller helps to reduce the difference between actual error rate and the reference error rate $(\dot{e} - \dot{e}_{NCT})$, especially near the origin.

As the results, the e is expected to decrease. Notice that, the structure of the AR-CM NCTF controller becomes a slightly complex with the additional controller elements; however, the design procedure remains easy and practical.

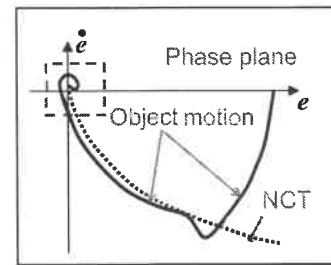


FIGURE 2. Problem occurred with the Continuous Motion NCTF controller.

Design Procedure

The design procedure of the AR-CM NCTF controller includes that of the CM NCTF controller. The mechanism with friction will be used in evaluating the usefulness of the

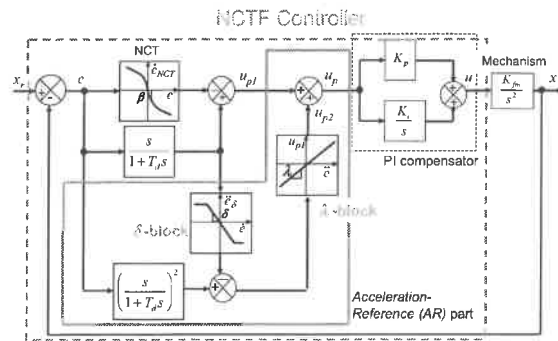


FIGURE 3. Acceleration Reference – Continuous Motion (AR-CM) NCTF control system.

proposed controller. In [10], it has been shown that the design procedure of the CM NCTF controllers is independent of friction. The constructed NCT and PI compensator in the CM NCTF controller for mechanism with friction [10] is used in implementing the AR-CM NCTF controller.

The AR-CM NCTF control system is basically designed according the following procedure.

- (I) *NCT construction & PI compensator determination*: The same constructed-NCT and PI compensator as the previous CM NCTF controller [10] is used.
- (II) *δ -block construction*: Constructed from the open-loop acceleration. Its inclination near origin, δ is the same as NCT, β .
- (III) *First-order low-pass filter design*: The first-order low-pass filter is used to reduce the gain at high frequencies. The selected f_d ($f_d=1/T_d$) is twice larger than the f_β ($f_\beta=1/2\beta$, which is the frequency determined from the inclination of NCT near origin). The controller includes many derivative elements which are responsible for the amplification of the measurement noise. The low-pass filters can reduce the bad influence of the derivative action. The selected time constant, T_d is useful to filter suitably the noise and to avoid the significant influence on the dominant dynamics of the control system.
- (IV) *λ -block construction*: The parameter λ is an adjustable gain. The λ gain is adjusted to reduce the residual vibration of the positioning performance. The λ might be any value of gain parameter.

It should be noted that the design procedure shown above can be completed without any motion equations, known model parameters and the basic control theory. Even though the complexity of the AR-CM NCTF controller is slightly increased with the additional controller elements, its design procedure remains easy and practical.

EXPERIMENTAL SETUP

The 1-DOF air-slide mechanism depicted in Fig. 4 is used to evaluate the design procedure of the AR-CM NCTF controller. The mechanism is fundamentally under non-contact condition. The adjusting unit for friction characteristic as shown is used when a mechanism with friction is needed. The contact condition between the plastic bar in the unit and the table is adjustable. In this paper, the mechanism with friction characteristic has friction coefficient, f_{max} 0.35N and damping coefficient generated by the grease is 9.3Ns/m.

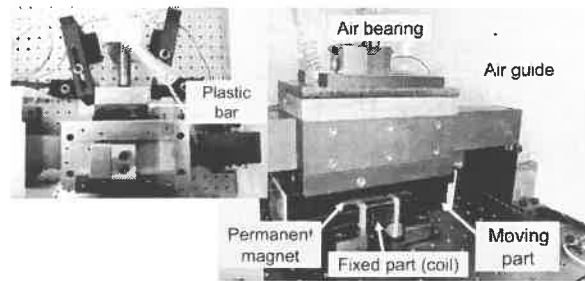


FIGURE 4. 1-DOF air-slide mechanism with the adjusting unit for friction characteristics.

PERFORMANCE EVALUATION

In order to show the effectiveness of the AR-CM NCTF controller, the PTP and tracking performances are experimentally evaluated. Then, its performances are compared with the CM NCTF controller. The AR-CM NCTF controller has the same controller parameters as the CM NCTF controller that shown in Table 1. Besides, the AR-CM NCTF controller is compared with the designed PD controllers with disturbance observers, in the purpose to examine and compare their robustness.

TABLE 1. Controller parameters

	βs^{-1}	δs^{-1}	λs^{-1}	K_p	K_i
CM NCTF	110	-	-	0.25	10.12
AR-CM NCTF	β	β	$1/3\beta$	0.25	10.12

Comparison of positioning performance

Figure 5 shows the comparison of the experimental positioning responses to 5mm step inputs with the AR-CM and CM NCTF controllers. The results obviously show that the AR-CM NCTF controller is able to reduce the undesired overshoot. The object motion of the AR-CM NCTF controller is able to follow the NCT motion accurately during following phase on the phase plane. This phenomenon is important in improving the transient response of the control system.

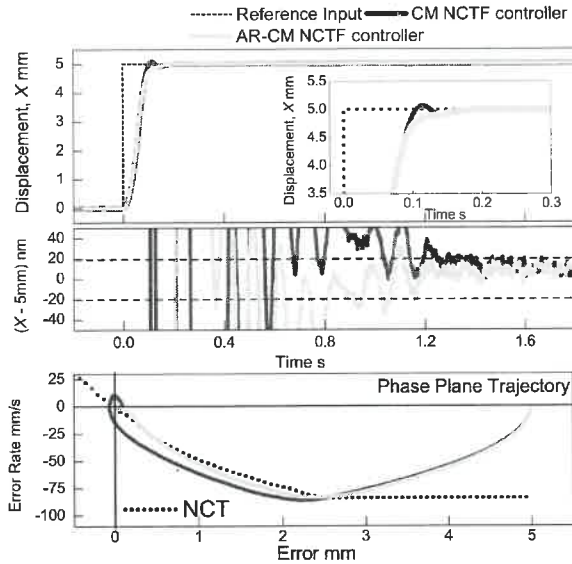


FIGURE 5. Step responses to a step input 5mm.

Comparison of tracking performance

Figure 6 illustrates the experimental tracking performance of the AR-CM and CM NCTF controllers. The experiments were conducted with a sine wave input which has frequency 0.5Hz with amplitude 1mm. Clearly, the AR-CM NCTF controller produces better tracking performance in smaller tracking error than the CM NCTF controller.

Comparison of the AR-CM NCTF controller and PD control with a disturbance observer

The disturbance observer controller offers several attractive features. It is relatively easier to design than the other advanced controllers and often shows higher robust to the disturbance force and model parameters, in comparison to the conventional PID controllers. Thus, in recent years, the disturbance observer has become one of the most commonly used schemes in industrial applications although it

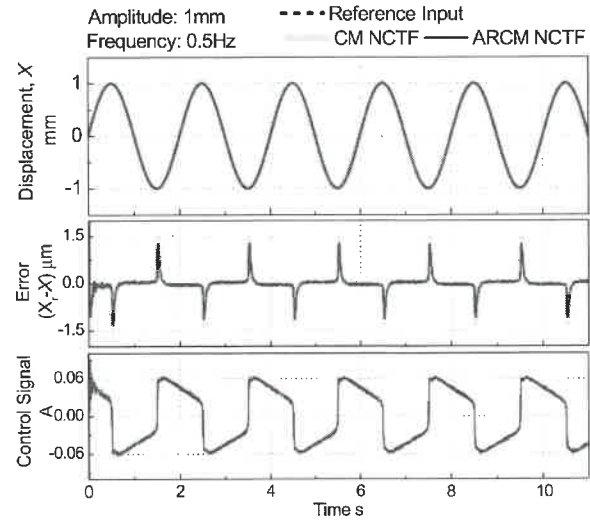


FIGURE 6. Tracking motion performance to a sine wave input 0.5Hz, 1mm.

takes more control theory to design it than the PID controller.

Design of PD controller with a disturbance observer (PDDO)

For comparison with the AR-CM NCTF controller, the PDDO controller is designed under the condition:

- (I) The linearized control system has the same bandwidth as the linearized AR-CM NCTF control system.

The state-equation of the object is derived as:

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dot{d}_{f_n} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\alpha & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ d_{f_n} \end{bmatrix} + \begin{bmatrix} 0 \\ K_{f_m} \\ 0 \end{bmatrix} u \quad .. (1)$$

where $d_{f_n} = \frac{1}{M}d_f$ and $K_{f_m} = K_f/M$ with d_f is disturbance force, K_f is force constant and M is total mass of the object. The d_{f_n}/K_{f_m} has the same unit as the control input.

The disturbance observer based on (1) can be designed to estimate d_{f_n} . As the disturbance observer, the minimum order observer [14] was designed. The PDDO control system is shown in Fig. 7. In the design of PDDO, the PD compensator parameters are determined so that the control system satisfies the condition (I).

Table 2 shows the controller parameters for two kinds of the designed PDDO. Both PDDO controllers have the same bandwidth as the AR-CM NCTF controller, however, the PDDO-B is designed to have exactly same frequency response as the AR-CM NCTF control system.

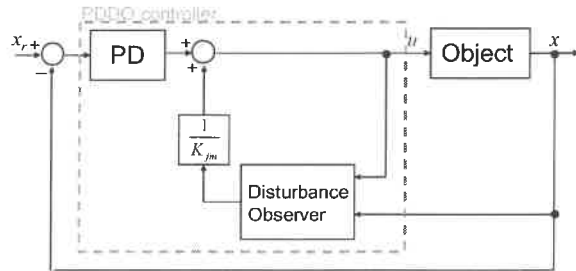


FIGURE 7. PDDO control system

TABLE 2. Controller parameters of PDDO control system

Controller	K_{op}	K_{od}	Observer's poles
PDDO-A	20.57	0.6	[0.1, 0.11]
PDDO-B	70.57	0.5	[0.1, 0.11]

Comparison of positioning performance

Figure 8 depicts the positioning responses of three controllers to a 1mm step input. The response with PDDO-B controller (which has the same frequency response as the AR-CM NCTF controller) produces overshoot. The PDDO-A and AR-CM NCTF controllers are able to eliminate overshoot, however, the PDDO-A controller shows slightly slower rise time than the AR-CM NCTF controller.

The positioning responses with increased mass (twice times) for three controllers are shown in Fig. 9. The PDDO-B still the one produces the highest overshoot among the three controllers. The AR-CM NCTF and the PDDO-A controllers maintain the no overshoot response. Overall, the AR-CM NCTF controller exhibits high positioning responses despite the changes of model parameters, than the both PDDO controllers.

Comparison of tracking performance

Figure 10 shows the tracking motion performances of the three controllers to a sine wave input (0.5Hz, 1mm). Even though the PDDO-A is able to demonstrate better PTP positioning performance, however, it yields extremely higher maximum tracking error in

tracking motion control. The PDDO-B, which has similar frequency responses as the AR-CM NCTF controller has shown the similar tracking responses as the AR-CM NCTF controller, with smaller tracking error. Nonetheless, the tracking error of the PDDO-B is increased (approximately 25%) when the mass of the object is increased (as shown in Fig. 11).

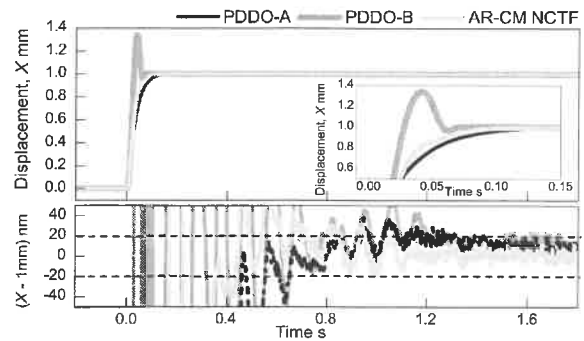


FIGURE 8. Step responses to a step input 1mm.

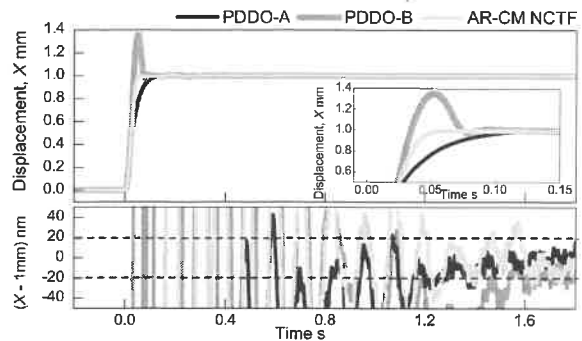


FIGURE 9. Step responses to a step input 1mm (increased mass object).

CONCLUSION

In the brief, the AR-CM NCTF controller was proposed as a new design approach to improve the motion control performance instead of the existing CM NCTF controller. The usefulness of the proposed controller was validated using mechanism with friction. The design procedure remains easy and practical. The experimental results proved that, the AR-CM NCTF controller is capable in reducing overshoot and motion error in comparison to the CM NCTF controller. Then, the positioning and tracking motion performances of the AR-CM NCTF controller was compared with the two PDDO controllers and evaluated by experiment. The AR-CM NCTF controller demonstrates the superior

positioning and tracking performances and has higher robustness to the mass increase than the two PDDO controllers. Overall, the AR-CM NCTF controller was proved to show the overshoot reduction characteristics and high disturbance reduction characteristics.

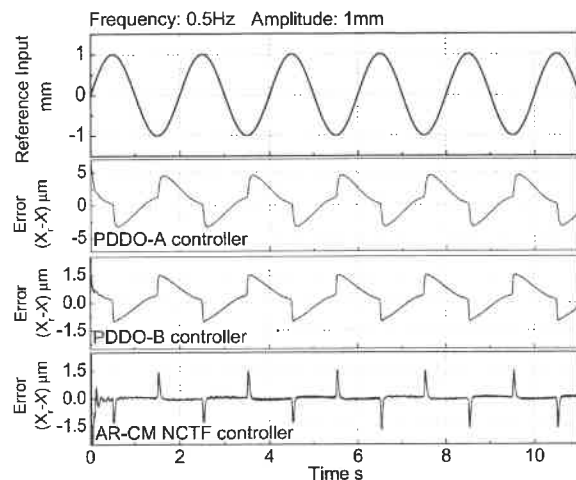


FIGURE 10. Tracking motion performance to a sine wave 0.5Hz, 1mm.

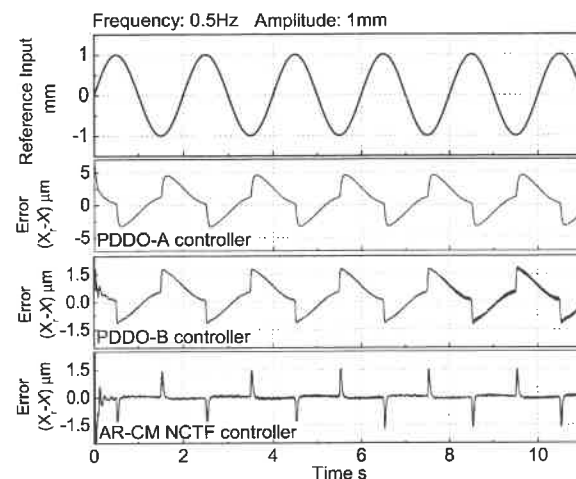


FIGURE 11. Tracking motion performance to a sine wave input 0.5Hz, 1mm (increased mass object).

REFERENCES

- [1] Cheng CC, Chen CY. A PID approach to suppressing stick-slip in the positioning of transmission mechanisms. *Control Engineering Practice*. 1998; 6:471-9.
- [2] Tan KK, Lee HL, Dou HF, Chin SJ, Zhao S. Precision motion control with disturbance observer for pulsewidth-modulated-driven permanent magnet linear motors. *IEEE Trans on Magnetic*. 2003; 39(3):1813-8.
- [3] Su YX, Zheng CH, Duan BY. Automatic disturbance rejection controller for precision motion control of permanent-magnet synchronous motors. *IEEE Trans on Industrial Electronics*. 2005; 52(3):814-23.
- [4] Yao B, Al-Majed M, Tomizuka M. High performance robust motion control of machine tools: An adaptive robust control approach and comparative experiments. *IEEE/ASME Trans on Mechatronics*. 1997; 2(2):63-76.
- [5] Zhu HA, Hong GS, Teo CL, Poo AN. Internal model control with enhanced robustness. *Int. J. Systems Sci.* 1995; 26(2):277-93.
- [6] Li YF, Wikander J. Model reference discrete-time sliding mode control of linear servo motor precision servo systems. *Mechatronics*. 2004; 14:835-51.
- [7] Wahyudi, Sato K, Shimokohbe A. Characteristics of practical control for point-to-point (PTP) positioning systems effect of design parameters and actuator saturation on positioning performance. *Prec Eng*. 2003; 27(2):157-69.
- [8] Maeda G.J, Sato K. Practical control method for ultra-precision positioning using ballscrew mechanism. *Prec Eng*. 2008; 32(4):309-18.
- [9] Sato K, Maeda G.J. A practical control method for precision motion – improvement of NCTF control method for continuous motion control. *Prec Eng*. 2009; 33(2):175-86.
- [10] Chong SH, Sato K. Practical controller design for precision positioning, independent of friction characteristic. *Prec Eng*. 2010; 34(2): 286-300.
- [11] Sato K, Maeda GJ. Practical ultraprecision positioning of a ball screw mechanism. *International Journal of Precision Engineering and Manufacturing*. 2008; 9(2):44-9.
- [12] Sato K, Maeda GJ. Fast precision positioning of a ball screw mechanism based on practical NCTF control. *Int. J. of Automation Technology*. 2009; 3(3):233-40.
- [13] Sato K, Nakamoto K, Shimokohbe A. Practical control of precision positioning mechanism with friction. *Prec Eng*. 2004; 28:426-34.
- [14] Ogata K. *Modern Control Engineering*. 4th ed. Prentice Hall: 2002. p. 871-81.

FREQUENCY DOMAIN BASED FEED FORWARD TUNING FOR FRICTION COMPENSATION

David J. Rijlaarsdam^{1,2}, Vincent van Geffen¹, Pieter W.J.M. Nuij¹
Johan Schoukens² and Maarten Steinbuch¹

¹Dept. of Mechanical Engineering / Control Systems Technology*
Eindhoven University of Technology
Eindhoven, The Netherlands

² Departement ELEC
Vrije Universiteit Brussel
Brussels, Belgium

ABSTRACT

In high precision motion control, performance is often limited by the presence of nonlinearities. In this study, the presence of nonlinear influences in a high precision transmission electron microscope stage is investigated using broadband multisine signals. These measurements yield the nature and level of nonlinearities as well as the best linear approximation of the dynamics. By quantitatively measuring the level of nonlinear influences, this method indicates the relevance of improved modeling. Next, the nonlinear influences are modeled explicitly by measuring the higher order sinusoidal input describing functions (HOSIDF) of the system which describe the 'direct' response of the system at the input frequency as well as at harmonics of the input frequency. Application of this technique yields a structured way to design Coulomb friction feed forward in the presence of nonlinearities. This procedure linearizes the input-output dynamics by applying feed forward and measuring the HOSIDFs which indicate the remaining nonlinear effects. Application of this technique yields a structured way to design feed forward in the presence of nonlinearities.

INTRODUCTION

When identifying and controlling (mechanical) systems, a linear model structure is often assumed. If nonlinear influences are small, such assumptions may be justified. In order to draw conclusions about nonlinear influences (type and magnitude) the authors in [1, 2, 3] present a multisine based, frequency domain identification

approach. This method yields both the Best Linear Approximation (BLA) of the systems dynamics and the magnitude and type of nonlinearities present. To further quantify the nonlinear influences the authors in [4, 5, 6] present a nonparametric modeling technique referred to as Higher Order Sinusoidal Input Describing Functions (HOSIDF). This technique describes the response of a system by relating the magnitude and phase of the harmonics of a sinusoidal input, in the output signal due to nonlinear influences. Finally, this study extends the concept behind the HOSIDFs to optimal feed forward design for a friction dominated motion system.

In the first section two experimental set-ups used in this study are introduced. Next, the multisine based approach is used to measure the BLA of the dynamics of an industrial high precision motion stage. This method emphasizes the relevance of improved modeling by **detecting** the magnitude and type of nonlinearities present. Moreover, the HOSIDFs of the same system are measured to further **quantify** the nonlinear behaviour. Finally, the concept behind the HOSIDFs is used to design optimal feed forward **control** for a 4th order motion system with Coulomb friction, minimizing the amount of nonlinear influence and largely improving the systems performance.

EXPERIMENTAL SET-UP

In this paper two experimental set-ups are considered. First, an industrial motion stage used to control the sample position in a transmission electron microscope (TEM) is used to demonstrate the different measurement techniques. This application is selected since the performance required in this application is particularly high. The high reproducibility and resolution required impose strong requirements on the control loop

*This work is carried out as part of the Condor project, a project under the supervision of the Embedded Systems Institute (ESI) and with FEI company as the industrial partner. This project is partially supported by the Dutch Ministry of Economic Affairs under the BSIK program. Corresponding author D.J. Rijlaarsdam, d.j.rijlaarsdam@tue.nl

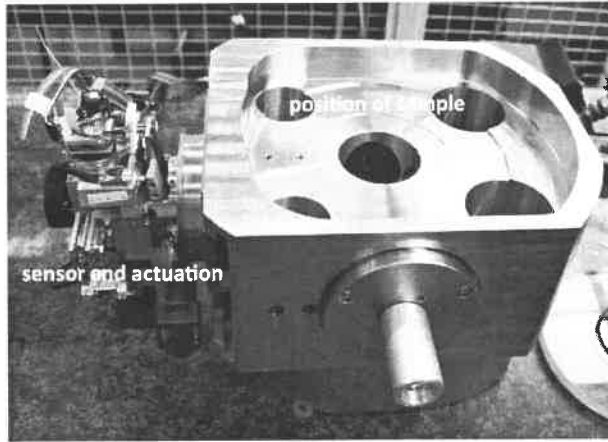


FIGURE 1. Experimental set-up (I): High precision TEM motion stage.

used to control the system. Second, a 4th order mass, spring, damper system with Coulomb friction is used to illustrate the usage of the measured nonlinear response in feed forward tuning. In the sequel the experimental set-ups are discussed in detail.

(I) Transmission electron microscope stage

Figure 1 depicts the TEM stage, removed from the TEM. The stage is used as a SISO system driven by a Maxon DC motor, with the engine voltage as input and the position of the stage as an output. The application of the excitation signal and measurement of the response is performed using a SigLab 20-42 dynamics analyzer providing 90 dB aliasing protection. The sample position is measured using a MicroE Mercury 3500 encoder with a sensor accuracy of 5 nm.

(II) 4th order mechanical system with friction

Figure 2 shows a 4th order mass, spring, damper system with Coulomb friction applied at the motor side. The system consists of two rotating masses connected by a torsional element. The friction is applied by applying a constant normal force acting through a vertical guidance system. The contact between the rotating mass and the friction element consists of two cylindrical elements resulting in an approximate single point contact. The constant normal force, combined with the single point of steel to steel contact results in a close approximation of Coulomb friction. The rotation of both masses is measured using an optical encoder with an accuracy of 500 increments per rotation and the system is driven by a Maxon DC motor. The engine voltage is the input of the sys-

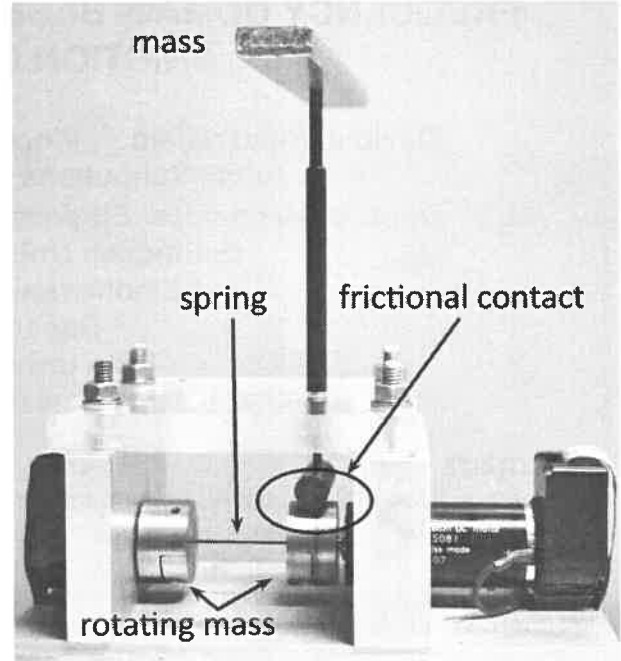


FIGURE 2. Experimental set-up (II): 4th order system with Coulomb friction.

tem, while the rotation of the mass on the motor side (right side in Figure 2) is the output.

DETECTING NONLINEARITIES

Methodology

In this section a multisine based method is introduced that allows measurements of both the best linear approximation of a system under test and the magnitude and type of nonlinearities present [1, 2, 3]. In general, output spectra Y are composed of the harmonic content generated from the input U by the best linear approximation $Y_{BLA} = H_{BLA}U$, stochastic disturbances (noise) Y_{noise} and disturbances generated by nonlinear influences Y_S [1]:

$$Y = Y_{BLA} + Y_S + Y_{noise} \quad (1)$$

Using signals with specific spectra and phase distributions called random odd multisines, the amount of information about nonlinear influences obtained from measurements is maximized. The m^{th} realization of a random odd multisine is defined as:

$$u^{[m]}(t) = \sum_{n=1}^N \alpha_n \sin(2\pi f_n t + \varphi_n^{[m]}) \quad (2)$$

with $\alpha_n \in \mathbb{R}_{\geq 0}$ possibly different for various frequencies, but constant over different real-

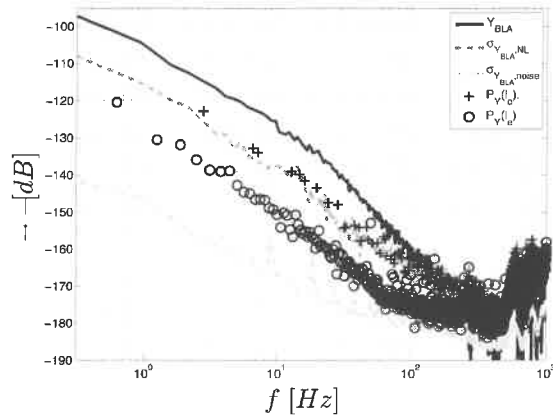


FIGURE 3. Output spectrum of a typical multisine experiment, showing the effects of nonlinearities ($\sigma_{Y_{BLA},NL}$) and noise ($\sigma_{Y_{BLA},noise}$) as variances on the best linear approximation of the spectrum ($u_{rms} = 3.72$ V).

izations. The phases $\varphi_n^{[m]}$ are randomly selected from the interval $[-\pi, \pi)$ for each realization and the odd harmonic frequency lines $f_n \in \mathbb{F}_o = \{(2k-1)f_0^b, k \in \mathbb{N}_1, f_0^b \in \mathbb{R}_{>0}\}$ are selected identically for all realizations. The random odd multisine is particularly useful for detection of nonlinear effects when a frequency line is removed randomly every 5 odd frequency lines [2]. For convenience the same spectral lines are removed for all realizations.

When using odd random multisines, there are two ways to detect nonlinearities in the spectral representation of the measured response. First of all, energy may appear on non-excited lines in the output spectrum. Secondly, a variance, larger than to be expected based on stochastic distortions, is observed on the measured output spectrum when using different realizations of the multisine. To detect energy at non excited frequency lines, a single measurement suffices classifying the type of nonlinear effects since energy may appear on non-excited odd or even lines. However, to obtain the variance on the output spectrum at excited lines due to noise and nonlinearities separately, multiple experiments with different realizations of the excitation signal are required. In this paper both methods are combined to both quantify the extend of nonlinear effects and classify them.

First, consider the excited spectral lines in the input signal. Performing M experiments with M re-

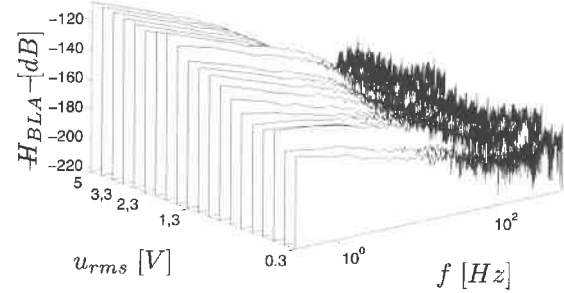


FIGURE 4. Best linear approximation H_{BLA} of the systems transfer function measured using the broadband excitation approach, as a function of input amplitude and frequency.

alizations of P periods the excitation signal yields $M \times P$ input and output time series and their corresponding spectra $U(\omega)$ and $Y(\omega)^{[m]}$. Averaging over multiple periods of the same realization yields the average spectrum and the variance on this spectrum due to stochastic distortions $\sigma_{Y,noise}^{2[m]}$, but not that due to nonlinear effects. Next, calculating the average spectrum over different realizations yields the best linear approximation Y_{BLA} of the true spectrum and the variance on this averaged spectrum due to nonlinearities $\sigma_{Y_{BLA},NL}^2$. Second, consider the response at odd (o) and even (e) non-excited lines $\mathcal{P}(\ell_{o/e})$. The spectrum has random phase at these lines, hence calculating the variance over the different measured spectra yields the average power that occurs at these frequencies. This yields a measure for nonlinear behaviour as well.

Application

The BLA and the level of nonlinearities in the industrial high precision stage are measured using the described procedure. Measurements are performed with measurement frequency of $f_s = 2560$ Hz and a block length of $N_{block} = 8192$ measurement points. This yields in a base frequency of the random odd multisine of $f_0^b = \frac{f_s}{N_{block}} = 0.3125$ Hz. Sufficient waiting time is allowed to assure that transient phenomena have damped out, avoiding leakage phenomena and no windowing is applied.

$M = 10$ realizations of the odd random multisine have been generated and the response has been measured for $P = 10$ periods. Furthermore, this experiment is repeated for 20 different rms values

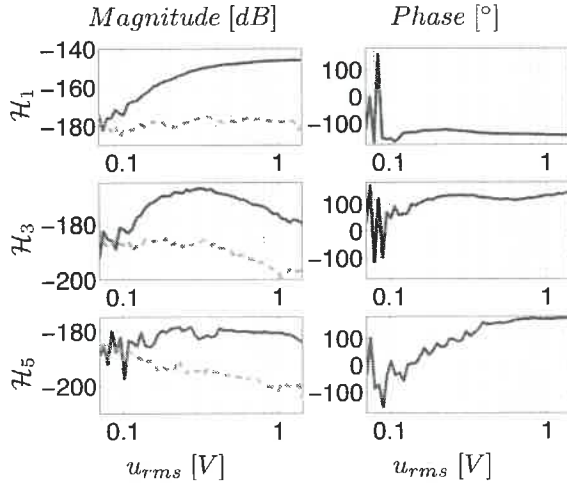


FIGURE 5. HOSIDFs at 20 Hz calculated from 10 experiment per amplitude. Mean value ($\bar{\mathcal{H}}_k$) of the HOSIDF (—) and standard deviation $s_{\bar{\mathcal{H}}_k}$ (---).

of the multisines, logarithmically scaled between 0.3 V and 5.0 V. A typical output spectrum is depicted in Figure 3. The best linear approximation of the systems dynamics is depicted in Figure 4 as a function of both frequency and input power.

Results

From Figure 3 it becomes clear that nonlinearities have an average level 10 dB lower than the power generated in the output spectrum by the BLA of the system. Both odd and even nonlinearities are detected, but odd nonlinearities dominate by almost 20 dB. Finally, the variation due to stochastic influences is almost 30 dB lower than that due to nonlinear effects.

QUANTIFICATION OF NONLINEAR EFFECTS

Methodology

In order to use the obtained information about nonlinearities from the previous section in controller design, the Higher Order Sinusoidal Input Describing Functions (HOSIDF) of this system are identified. HOSIDFs describe not only the 'direct' response (gain and phase) of a system at the excitation frequency, but describe the response at harmonics of the excitation frequency as well. Consider the following input signal used to identify the HOSIDFs:

$$w(t) = \beta \sin(2\pi f_0^s t) \quad (3)$$

Next, the k^{th} order HOSIDF is defined as:

$$\mathcal{H}_k(\beta, f_0^s) = \frac{Y(k f_0^s)}{U(f_0^s)} \quad (4)$$

$$\angle \mathcal{H}_k(\beta, f_0^s) = \angle Y(k f_0^s) - k \angle U(f_0^s) \quad (5)$$

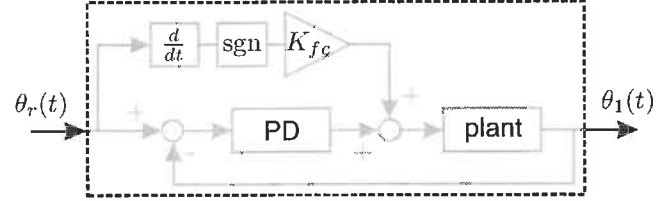


FIGURE 6. Control loop.

where, $Y(k f_0^s) \in \mathbb{C}$ is the output spectrum at the k^{th} harmonic frequency line and $U(f_0^s) \in \mathbb{C}$ the spectral content of the input signal at its fundamental frequency $f_0^s \in \mathbb{R}_{>0}$. HOSIDFs are generally a function of both the input frequency f_0^s and amplitude $\beta \in \mathbb{R}_{>0}$. To assess the quality of the estimates of the HOSIDFs, multiple experiments are conducted and the average of the k^{th} HOSIDF is defined as $\bar{\mathcal{H}}_k$ and its variance as $s_{\bar{\mathcal{H}}_k}^2$.

Application

The TEM motion stage was excited with frequencies ranging from 5 Hz to 300 Hz in steps of 5 Hz. Each response has been measured 10 times for input powers ranging from 0.07 V to 1.41 V (rms). All measurements have been performed using a SigLab measurement system with a measurement frequency of 5120 Hz and a block length of 2048 points. This results in leakage free measurements since sufficient waiting time is allowed for the transient response to damp out.

Results

A typical series of HOSIDFs is depicted in Figure 5. Odd HOSIDFs are considered, since the even HOSIDFs are very low. The system becomes more linear for increasing value of u_{rms} (decreasing gradient $\partial \mathcal{H}_1 / \partial u_{rms}$ and decreasing $\mathcal{H}_i$ $i > 1$). The HOSIDFs provide other information about the nonlinear behaviour as well, such as the maximum increase in nonlinear behaviour that can be related to stick-slip transfer [6] and a maximum in nonlinear influences. However, these are outside the scope of this paper.

OPTIMAL FEED FORWARD DESIGN

It was shown that spectral information in the output spectrum can be used to assess the influence of nonlinear phenomena. The next issue addressed in this paper is how to use this information in control, by tuning a feed forward controller based on frequency domain information.

Set-up

The system in Figure 2 is used in feedback with

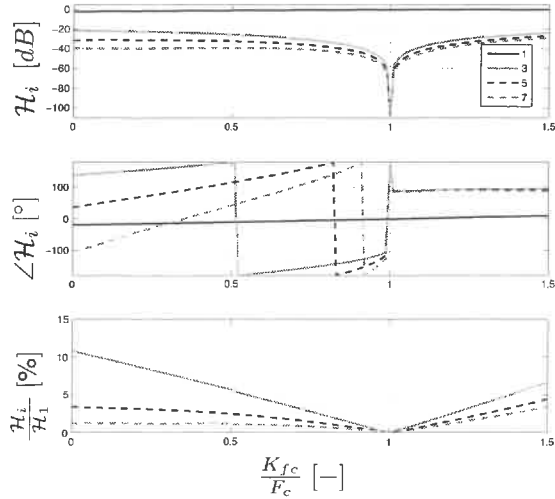


FIGURE 7. Simulation results. (top, center) Magnitude and phase of the first three odd HOSIDFs. (bottom) Ratio between odd nonlinear and linear response components.

a PD-controller and a reference signal θ_r as depicted in Figure 6. The rotation of the motor $\theta_1(t)$ is the output and controlled variable. Apart from feedback, a nonlinear feed forward is applied to compensate for Coulomb friction. Using a stabilizing PD controller, the feed forward parameter K_{fc} will be tuned to linearize the input-output behaviour of the closed loop system. In other words, K_{fc} will be tuned such that the amount of nonlinear (harmonic) content relative to the linear content in the output is minimal, or in terms of the systems HOSIDFs:

$$K_{fc}^* = \underset{K_{fc} \in \mathbb{R}_{\geq 0}}{\operatorname{argmin}} \frac{\mathcal{H}_i(K_{fc})}{\mathcal{H}_1(K_{fc})} \quad (6)$$

In the sequel, numerical and experimental results are provided that illustrate this tuning method. In both simulations and experiments K_{fc} is varied from 0 (no feed forward) until the system is slightly overcompensated. The HOSIDFs are measured yielding the required minimum. Since the feed forward friction model is a static model, the tuning procedure is performed for one frequency only, assuming Coulomb friction in the plant.

Numerical Results

Figure 7 depicts simulation results of a two mass, spring, damper system similar to the system depicted in Figure 2. The system is subject to Coulomb friction at the motor side and operates in feedback as depicted in Figure 6. From Figure 7 it becomes clear that input-output behaviour of

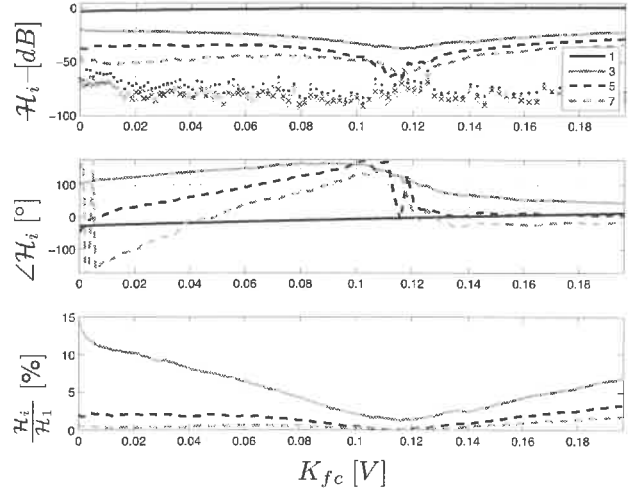


FIGURE 8. Measurement results. (top, center) Magnitude and phase of the first three odd HOSIDFs. (bottom) Ratio between odd nonlinear and linear response components. (— — — average, · × standard deviation)

the system becomes more linear with increasing K_{fc} until an optimum is reached when the feed forward equals the Coulomb friction force. Furthermore, the phases of the HOSIDFs turn 180° at the optimal setting.

Experimental Results

Figure 8 shows the same behaviour in experiments using the experimental set-up depicted in Figure 2. An optimum is reached at $K_{fc}^* = 0.1157$ V where the relative level of nonlinearities has decreased from 15% to less than 1.5%. The remaining nonlinear influences are due to effects that are not captured by the feed forward model. Note that even nonlinearities (not depicted) have a relative level of only 4% and are not influenced by the purely odd feed forward.

Discussion

The method presented in this paper enables optimal tuning of (feed forward) parameters in the sense that the input-output behaviour of the system is linearized. The procedure has been demonstrated for Coulomb friction feed forward but may be used to tune arbitrary controllers in the presence of nonlinearities as long as the complete system has a periodic response to a periodic input.

The optimal value K_{fc}^* not necessarily yields the smallest tracking error. Figure 9 shows the response and error observed in experiments: in ab-

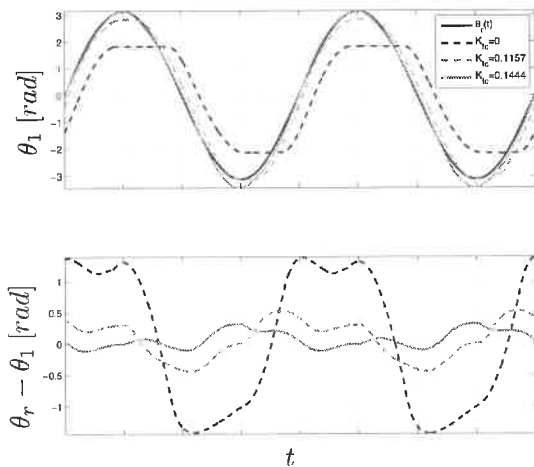


FIGURE 9. Closed loop response with no feed forward ($K_{fc} = 0$), optimal linear response ($K_{fc}^* = 0.1157$), minimal rms error ($K_{fc} = 1444$). (top) Time response. (bottom) Error.

sence of feed forward ($K_{fc} = 0$), with optimal feed forward ($K_{fc} = K_{fc}^* = 0.1157$) and at the smallest observed error ($K_{fc} = 0.1444$). It appears that by increasing the feed forward by 25%, the response is still improved. However, for $K_{fc} > K_{fc}^*$ the feed forward compensates apples with oranges. At $K_{fc} = K_{fc}^*$ the feedforward compensates nonlinear influences optimally for the chosen feed forward model. Remaining (non)linear error sources are not captured by the feed forward model. For higher K_{fc} the Coulomb feed forward starts to compensate other effects. In this case, the feed forward most likely compensates viscous damping, yielding a local optimum valid only for the type and level of excitation used. Instead, K_{fc}^* should be selected and a velocity proportional feed forward added. Furthermore, the time varying nature of friction due to wear and tear becomes apparent during longer experiments requiring an adaptive model structure for long term compensation.

CONCLUSIONS AND FUTURE RESEARCH

Two methods that allow detection and quantification of nonlinear effects are introduced and used to identify nonlinear influences in an industrial high precision motion stage. Multisine based measurements yield the nature and level of nonlinearities as well as the best linear approximation of the dynamics. Next, nonlinear influences are modeled by measuring the HOSIDFs of the system under test. Finally, the concept behind the

HOSIDFs is used to optimally tune a feed forward controller.

The presented method allows tuning of arbitrary controllers linearizing input-output dynamics of systems with a periodic response to a periodic input. It may be used to assess and compare the quality of different controllers and future research aims at optimization algorithms for fast, automated controller tuning. Furthermore, tuning of multiple parameter (non)linear and adaptive controllers is investigated.

ACKNOWLEDGMENTS

The authors thank Bas van Loon for the measurements of the HOSIDFs and Johan Pattyn for his contribution to the realization of the set-up (II).

REFERENCES

- [1] Pintelon, R., J. Schoukens (2001). *System identification: a frequency domain approach*, IEEE Press Piscataway, NJ, (ISBN 0-7803-6000-1).
- [2] Pintelon, R., G. Vandersteen, L. De Locht, Y. Rolain, J. Schoukens (2004). Experimental characterization of operational amplifiers: a system identification approach, *IEEE Trans. Meas. Instr.*, vol. 53, no. 3, pp. 854–867.
- [3] Vanhoenacker, K., T. Dobrowiecki, J. Schoukens (2001). Design of multisine excitations to characterize the nonlinear distortions during FRF measurements, *IEEE Trans. Meas. Instr.*, vol. 50, no. 5, pp. 1097–1102.
- [4] Nuij, P.W.J.M., O.H. Bosgra, M. Steinbuch (2006). Higher-order sinusoidal input describing functions for the analysis of nonlinear systems with harmonic responses, *Mechanical Systems and Signal Processing*, vol. 20, pp. 1833–1904.
- [5] Nuij, P.W.J.M., M. Steinbuch, O.H. Bosgra (2008). Measuring the higher order sinusoidal input describing functions of a nonlinear plant operating in feedback, *Control Engineering in Practice*, vol. 16, pp. 101–113.
- [6] Nuij, P.W.J.M., M. Steinbuch, O.H. Bosgra (2008). Experimental characterization of the stick/sliding transition in a precision mechanical system using the Third Order Sinusoidal Input Describing Function, *Mechatronics*, 18(2), pp 100-110.

A HIGH PERFORMANCE X-Y STAGE WITH A NOVEL TOPOLOGY

Kevin J. McCarthy¹ and Bradley L. Hunter²

¹Dover

Westboro, MA, USA

²Active Precision

Burlington, MA, USA

This paper describes a 350 mm travel X-Y positioning system with a novel topology, high dynamic performance, and an advanced real-time controller. While the initial implementation positions a 300 mm semiconductor wafer under a stationary laser system, the design can be scaled upwards or downwards in travel to suit a wide range of potential applications. The Delta Stage described below provides substantially higher performance than traditional X-Y stages, and its advantages become even more compelling as the travel is increased.

EXISTING SOLUTIONS

While traditional single axis stage designs can, with moderate effort, provide high dynamic performance, the design challenges and trade-offs grow substantially as the number of axes is increased. In particular, the requirement to position a payload in two axes with high dynamic performance encounters a number of design conflicts. The simplest expedient, that of stacking two similar axes, one above the other, suffers from excessive moving mass and low frequency structural resonances.

One reasonably efficient solution, referred to as a split-axis design, separates the X and Y axes, one of which carries the payload, with the second axis carrying the tool. While this approach is well suited to compact tools that bear acceleration well, in many cases the tool is prohibitively massive and/or includes sensitive components, and cannot be moved. In this case, the payload must be moved in two axes.

Several decades of evolution in X-Y stage design have converged to yield two related topologies for moving a payload in two axes, commonly referred to as "H" and "T" stages. In both cases, the payload carriage moves in a plane over a very flat surface, supported

vertically by a single preloaded planar air bearing. This bearing constrains movement of the carriage and payload in the Theta X, Theta Y, and Z axes, while providing freedom in the X, Y, and Theta-Z axes. A moving beam with a linear bearing and a linear motor constrains movement in the Y and Theta-Z axes, while controlling movement along the X axis. The distinction between the "H" and "T" design variants is that in the "T" design, this moving X axis beam is rigidly coupled to a bearing that travels along a stationary Y axis beam, and is driven by a single motor. In the "H" design, two motors are used to drive each end of the X axis beam along the Y axis, with a flexible joint between the X beam and the Y axis bearing. Feedback is provided by three linear encoders, supplemented in high accuracy designs by a laser interferometer and a pair of perpendicular flat mirrors mounted to the carriage assembly. For short travel variants, a single planar X-Y encoder can be mounted in the base of the carriage.

While commercially available from a number of suppliers, X-Y stages based on a "H" or "T" design suffer from some fundamental drawbacks, as listed below:

- The mass of the moving X axis beam and its linear motor limits performance.
- The moving X beam is subject to deflection and resonances due to bending moments.
- The payload is closely coupled thermally to the X axis motor coil.
- In "T" designs, cantilevered loads deflect the Y axis bearing.
- In "H" designs, the flexible coupling between the X beam and the Y bearing limits performance.

The above drawbacks manifest themselves in three ways: lower servo bandwidth (due to a low first resonance), limited acceleration (due to high moving mass and motor thermal limits), and thermal deformation of the payload (due to acceleration-induced heating). Attempts to mitigate these effects lead to extensive light-weighting and exotic materials, but for applications which are acceleration-bound, these efforts very rapidly encounter diminishing returns. The use of larger motors might seem helpful, but fails due to the added moving mass of the X axis motor stator. Motor thermal power is equal to $a^2 \cdot m^2 / K_m^2$, where a is the acceleration, m is the mass, and K_m is the motor constant (in Newtons per root watt). Since acceleration is what we want, and heat is its undesirable by-product, we can define a figure of merit K_s , equal to K_m^2 / m^2 , such that stage power equals a^2 / K_s . Motor thermal power is inversely proportional to K_s , and stage designs can be ranked on their value of K_s . Traditional "H" and "T" stages have K_s values that tend to range between 0.05 and 0.25.

THE DELTA STAGE

The Delta Stage (Fig. 1) was developed after attempts to raise the continuous acceleration capability and servo bandwidth on traditional stage designs had reached an impasse.

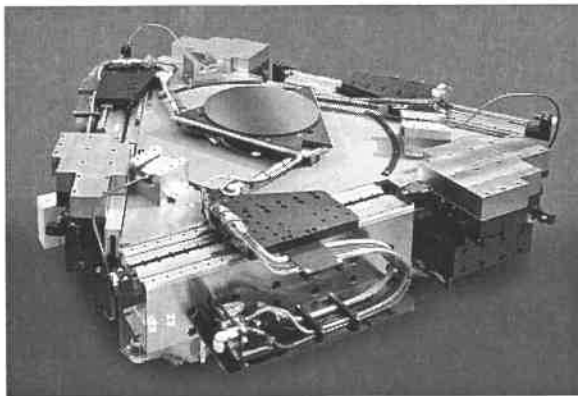


Figure 1. The Delta Stage

The Delta Stage utilizes a parallel link architecture, which has been known since the 1800's to be vastly superior to serial link mechanisms in agility (acceleration per watt) and stiffness [1]. "T" stages are serial link devices, while "H" stages employ a hybrid parallel/serial link design. The basic elements of the Delta Stage are a triangular base (granite in this case, although other materials can be used), three primary linear actuators (one on each side of the granite triangle), three counterforce linear

actuators, three tubular silicon carbide drive links, and a vacuum-preloaded central chuck-puck that carries the payload (in the first implementation, a 300 mm silicon wafer). The work area is a 650 mm circle, and the three drive links can rapidly position the chuck-puck anywhere within that circle. Three steel risers support the customer's optical plate, as well as mounting the fiber-optic interferometer heads.

Note that while parallel link mechanisms have a long history, the particular topology of the Delta Stage is unique, and several patents on the design are pending. Specific details and advantages of the Delta Stage are discussed below.

THREE DEGREES OF FREEDOM

The Delta Stage is a three DOF system, and positions the chuck-puck in three axes: X, Y, and Theta-Z. The ability to provide Theta-Z positioning with no additional penalty in moving mass, compromised dynamics, overall height, or cost is a distinct benefit in many applications. For the initial application, the chuck-puck rotates 20 degrees during wafer load/unload to minimize the reach required of the wafer handling robot. Future versions of the Delta Stage will redesign the current monolithic chuck-puck as a three DOF Z – Tip – Tilt assembly, providing precise positioning with high dynamics in all six degrees of freedom.

HIGH SPECIFIC STIFFNESS (Stiffness / Mass)

The three linear actuators drive the puck by applying tension and compression forces directly along the drive links. Since there are rotary joints at each end of the links, the links cannot be placed in bending. For a given link length, a tube in tension/compression can be made much stiffer with substantially less material (lower mass) than a beam that must both withstand bending forces and carry a large linear motor.

The time constant τ for settling after a move is inversely proportional to servo bandwidth, which is in turn proportional to the frequency of the first structural resonance. This first resonance scales as the square root of stiffness divided by mass, so high dynamic performance is a matter of maximizing the system's specific stiffness (stiffness/mass). The replacement of the large mass of the moving X beam and its heavy linear motor stator with a simple cylindrical drive link substantially decreases the moving mass, while the elimination of any bending moments substantially increases the stiffness of this drive

method. The result is a much higher ratio of stiffness to mass, and a considerably higher first resonance. In the first-article Delta Stage, a monolithic machined chuck-puck design was chosen, to reduce risk and lead time. With a chuck-puck moving mass of 4 kg, the first resonance was over 700 Hz, and the second resonance was over 1100 Hz. More aggressive light-weighting (e.g. expanded honeycomb + sheet metal) could achieve a first resonance over 1 kHz.

As stage travel and hence dimensions increase (as is happening in the current move from 300 mm to 450 mm wafers), the stiffness in bending of the moving X-axis beam of an "H" stage decreases with the cube of its length. The stiffness of the Delta Stage drive links, which experience only tension and compression, decrease linearly with length. The dynamic advantages of the Delta Stage, already substantial at 350 mm travel, become even more compelling when the travel grows to 500 mm or more.

HIGH AGILITY (Acceleration / watt)

The high agility of the Delta Stage results from two factors: low moving mass, coupled with large linear motors. High agility is important in reducing the power dissipation in the motors when performing high-duty cycle moves at high peak acceleration. For a point-to-point move of a prescribed distance and move time, the energy dissipated in the motors during that move will be inversely proportional to the previously defined figure of merit K_s .

The Delta Stage (like all parallel link mechanisms) has a dramatic advantage over conventional "H" stages due to the reduction in moving mass of the stage, together with the ability to use very large linear motors with only a modest increase in moving mass (since all three linear motor stators are stationary). In the Delta Stage, K_s varies a bit from place to place, but averages 12, compared to a value of 0.25 for a reasonably aggressive "H" stage. As a result, the Delta Stage is nearly 50 times more efficient, and **will dissipate ~2% of the power that would be dissipated by a conventional "H-Stage" performing the identical motion profile!** During 3G moves at 100% duty cycle, the Delta Stage motor coils dissipate less than 15 watts, and due to their large size, are barely warm to the touch.

THERMAL ISOLATION

In addition to substantially raising the continuous acceleration capability of the system, the very high efficiency of the Delta Stage reduces thermal coupling to sensitive system components and the payload, and substantially reduces their time and duty-cycle dependant thermal expansion and contraction. Secondary benefits include smaller amplifiers, smaller cables (less cable drag during stage motion), and smaller power supplies. The Delta Stage offers the further benefit of moving the motor thermal dissipation (what little there is) to the periphery of the stage base. In a conventional "H-Stage", at least one motor will dissipate power immediately under or adjacent to the payload. Given significant power dissipation in the immediate vicinity of the payload, conventional stages are challenged to provide both high precision and high accuracy.

FORCE CANCELLATION

In conventional "H" stages, the high moving mass produces large horizontal inertial forces during high acceleration, and the shift in CG during motion produces large and variable vertical forces. Most high accuracy systems require a vibration isolation system that reacts to these large inertial and CG-shift forces, significantly degrading the overall dynamic performance. While the Delta Stage moving mass is considerably lower than a conventional "H" stage to begin with, its three short travel counter-force actuators (only practical in a system with such high efficiency) cancel nearly all horizontal inertial forces. The control law for the counterforce actuators moves their reaction masses in opposition to stage motion, so as to maintain a perfectly fixed and central CG location for the entire X-Y stage sub-system. In many conventional implementations, the full acceleration capability of the stage must be pared back due to interactions with the isolation system; in the Delta Stage, operation at full acceleration imparts negligible forces to the isolation system.

INTEGRAL ACTIVE VIBRATION ISOLATION

Three highly sensitive geophones are built into the base of the Delta Stage. Their signals (resulting from external accelerations) are processed by the system controller, which in turn generates commands to the counterforce actuators that cancel those external accelerations. Unlike conventional active vibration cancellation systems, however,

vibration cancellation via the counterforce actuators imparts no forces to the frame, and cannot couple back into the isolation system. Typically, only passive vibration isolation is required with the Delta Stage, and due to the complete lack of CG shifts, active vertical isolation can be effected with the very simple and inexpensive addition (fully supported by the system controller) of three additional geophones and three small solenoids. The cost savings and reduction in overall complexity provided by the Delta Stage's integral, high-performance active vibration isolation system can be substantial.

BEARINGS AND MOTORS

The chuck-puck base bearing is a planar, low-noise, vacuum preloaded air bearing. The hinge and pivot joints at each end of the three drive links employ multiple sets of preloaded angular contact bearings. Three long recirculating ball modules with low lateral jitter guide the three primary actuators. Crossed roller ways guide the three counterforce actuators.

Three phase brushless linear servo motors were used to drive both the primary and counterforce linear actuators; K_m for the primary actuators was 23 Newtons per root watt. The motor coil mass is 37% of the total actuator moving mass, compared to 2-5% for traditional "H" stages.

FEEDBACK

Position feedback for each of the primary linear actuators consists of a pair of Heidenhain LIF linear encoders, which provide a sine/cosine output with a 4 micrometer electrical period. The system controller performs advanced interpolation on these signals, with a resolution of 78 picometers. One of the pair is located very close to the center of mass (and drive) of each actuator, while the second is located just below the hinge joint of the drive link. The use of two encoders allows the controller to measure and correct in real time any acceleration-induced yaw of the primary linear actuator. While this proved unnecessary in this first embodiment using mechanical bearings, this capability should be very valuable in subsequent air-bearing iterations of the primary linear actuators.

The exact position of the chuck-puck is determined in real time via closed-loop interferometer feedback, using a pair of plane mirrors integral to the monolithic chuck-puck (see below), and a stationary Renishaw RLE-10 two-axis, plane mirror interferometer, a

homodyne position sensor whose native output is sine/cosine. The interferometer's electrical signal period of 158 nanometers is interpolated by the system controller, resulting in a resolution of 3 picometers (well below its noise floor), and the Delta Stage's top speed is limited only by the interferometer, which is capable of operating at velocities of up to 1.0 meter/second. Thanks to the exceedingly real-time nature of the RDI based control system, laser firing pulses can be output directly from the interferometer position data, with a timing uncertainty of 1 nanosecond (1 nanometer at 1 meter per second).

INTEGRAL, DIAMOND-TURNED INTERFEROMETER MIRRORS

In most X-Y stages equipped with a plane mirror interferometer, two individual glass "stick mirrors", or a single "L mirror", are mounted to the upper axis carriage, and present two plane mirror surfaces at 90 degrees that reflect the two interferometer beams. Both of these approaches add substantial cost and mass to the system, and it is a difficult design challenge to clamp these mirrors securely without compromising either mirror flatness or system dynamics. In the Delta Stage, the two mirrors are oriented at 60 degrees, and have been directly diamond-turned into an electroless-nickel coating applied to lightweight ribbed extensions of the monolithic chuck-puck. The top of the mirror surface is 1.5 mm below the top of the wafer, and the mirrors are both diamond-turned and subsequently profiled via interferometry with an active chuck-puck base bearing, ensuring minimal deformation from flatness of the two surfaces in operation. Deviation from flatness is +/- 110 nm over full travel (350 mm), and the mirror first resonance is 1200 Hz. The total incremental moving mass of the integral interferometer mirrors is a mere 150 grams. The simple addition of a column reference mirror in the reference leg of the interferometer eliminates any sensitivity to atmospheric refractive index changes.

MONOLITHIC CHUCK-PUCK

The monolithic chuck-puck was machined in two halves from aluminum plate, producing a pair of structures with 1 mm hexagonal cell walls and dedicated compressed air and vacuum routing. These two halves were then vacuum braised together, producing a lightweight and stiff chuck-puck assembly.

Dedicated wafer chucks typically add significant cost, mass, and dynamic compromises to

traditional X-Y stage solutions. In the Delta Stage, an array of pins in the top of the monolithic chuck-puck forms an integral wafer chuck that provides minimal back-side contact, while ensuring rapid air flow to and from (and across) the vacuum chuck. The Delta Stage vacuum chuck adds no mass, no dynamic complications, and minimal incremental cost to the system, and consists of a large number of features photo-etched into the thin top skin of the monolithic chuck-puck.

Complementing the pin-grid wafer vacuum chuck is an integrated, 6 mm travel, three-post lift-pin assembly. The three moving posts have an internal switched vacuum to grip the wafer, and transport the wafer 6 mm above the vacuum chuck to provide access for a wafer handling robot paddle. The lift-pin assembly is driven by an integrated stepping motor linear actuator, and optical limit switches define the travel range and confirm lift-pin retraction. The total incremental mass of the lift-pin assembly is 75 grams.

PERFORMANCE

Figure 2 shows the performance of the Delta Stage while performing a typical semiconductor application. The stage is performing a series of operations "on-the-fly". The operations occur during brief moments when the stage is moving at constant velocity. The stage makes a cycloidal acceleration between the constant velocity segments. Note that the following error during 2.7 G acceleration segments is never more than a few microns, and settles to a few hundred nanometers in under 10 milliseconds after the acceleration transient.

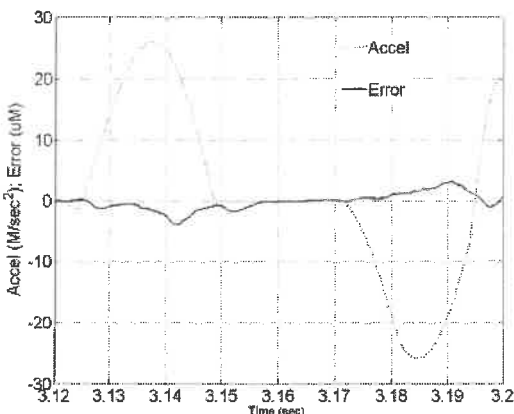


FIGURE 2. Stage acceleration and following error in the X direction during typical semiconductor processing.

DELTA STAGE CONTROLLER

The Delta Stage Controller is implemented with an infrastructure called RDI. RDI (Real-time Data Interface) provides synchronization and communication services in support of high-performance control systems. The Delta Stage Controller consists of a host computer where the application interface is implemented, an embedded real-time computer that executes the time-critical control code, and peripheral devices that interface with encoders, interferometers, amplifiers, galvos, and other application-specific devices.

The host and real-time computers each include an RDI PCI-card. Each card provides four (4) RDLINK ports that use Cat6 cables to communicate with other RDI-devices. Each port transmits a 100 MHz clock signal. The receiving port synchronizes to the RDI clock signal and adjusts the phase at the receiver to compensate for transmission delays. The resulting clock on the receiving port exhibits less than 0.1 nanosecond of jitter (3-sigma) and less than 1 nanosecond of skew relative to the clock source. This level of synchronization is critical to achieving nanometer-class precision when operating at speeds approaching 1 meter/sec.

The RDLINK data paths transmit data bi-directionally at 500 Mbit/sec. All connections are transformer isolated to eliminate ground loops. Each link can be up to 10 Meters in length. RDI allows components to be distributed throughout a system, without compromising timing precision. RDI extends the industry trend towards distributed motion control to the realm of nanometer precision and sub-nanosecond timing jitter.

The host and real-time computers transmit data over the RDLINK cable using real-time DMA transfers. Each RDI card includes 8 DMA engines that can be scheduled to transmit data packets at specific points in time relative to the servo period. Data from a peripheral card that is responsible for encoder or interferometer input is scheduled to arrive a few microseconds before the start of a servo period. The peripheral card transmits a packet of data directly into the memory of the real-time computer. When the servo period begins, the real-time computer finds the data packet in memory and begins calculations immediately. At the end of the servo calculations, the real-time processor writes output values to a fixed location in memory. A scheduled DMA transfer transmits

the output data to a peripheral connected to amplifiers. Note that since all data transfers are performed by the RDI hardware using DMA transfers, the real-time processor is free to focus exclusively on servo calculations. Typical packet transfers between servo-related peripherals and the real-time controller require less than 2 microseconds to complete.

The real-time computer in the Delta Stage Controller consists of an industry-standard PC mainboard with a Pentium 4 processor operating at 3.4 GHz. All PC peripherals (USB, LAN, RS232) and all interrupts are disabled in the real-time computer since the RDI hardware provides all communication services. The entire Delta Stage control program fits within the on-chip cache of the Pentium processor. With no other peripherals on the PCI bus, the RDI card is able to use the PCI bus exclusively for packet transfers. Since the Delta Stage Controller runs without using interrupts, the on-chip memory cache remains coherent between servo periods contributing to improved calculation performance. Typical performance exceeds 1 GigaFlop of double-precision operations per second.

The host computer sends commands to, and receives status from, the real-time computer using DMA transfers over the RDLink connection. In the Delta Stage Controller, the host application computes trajectories that the stage is to follow. Each entry in a trajectory specifies positions, velocities and accelerations as well as application-specific control settings corresponding to a single time-step of the servo. In some applications, the trajectories are computed and stored in a file before the actual motion occurs. In other applications, the trajectory must be computed while the stage is moving. In either case, the RDI hardware handles the transmission of trajectory packets from the host to the real-time computer without processor intervention on either side of the link. There is no practical limit to the duration of a trajectory that the Delta Stage can follow without stopping.

A state-space controller implemented using the RDI infrastructure controls the Delta Stage. The host processor specifies trajectories in a user-specified Cartesian coordinate system. The real-time processor performs all necessary transformations (geometric and kinematic) to transform the trajectory into a body-centered coordinate system.

The state-space controller closes the servo loop in the body coordinate system. The controller output is transformed into the actuator coordinate system where it is sent to the servo interface hardware in a DMA packet.

Sensor inputs are transformed from sensor coordinates to body coordinates and combined with body forces in the state observer to update body states. The state outputs are also used to project the vector of following error onto the nominal path of the stage. The component of error in the direction of motion is converted to an equivalent error in time. The time-based adjustment is used (in the Servo Interface FPGA) to correct nominal trigger times of signals that coordinate external devices with stage motion. Thus, the stage precision in the direction of motion for on-the-fly operations is limited only by the precision of the stage metrology (in this case, a laser interferometer), not by servo performance.

The trajectory planner executing on the host computer is responsible for deriving trajectories that are continuous through 3 derivatives (continuous acceleration, finite jerk) and within the performance capabilities of the stage (position, velocity, acceleration and voltage limits). The planner uses cycloidal motion profiles to link any vector position and velocity to any other vector position and velocity. Using acceleration feedforward from the trajectory, the stage typically will execute a 3G acceleration move with less than 2 microns of deviation from the trajectory during the acceleration phase. The stage settles within a few milliseconds after an acceleration transient to under 0.25 micron following error during a continuous velocity segment. Note that the timing of an external event with the stage motion is corrected for any following error, yielding noise-limited precision in the direction of stage motion.

REFERENCES

- [1] Merlet J.-P., *Parallel Robots*, 2nd Edition. Springer. Dordrecht, The Netherlands: 2006.

DESIGN AND CONTROL OF OVER-ACTUATED LIGHTWEIGHT 450 MM WAFER CHUCK

Dick A.H. Laro¹, Rein Boshuizen^{1,3}, and Jan van Eijk^{2,3}

¹MI-Partners BV

Eindhoven, the Netherlands

²MICE BV

Eindhoven, the Netherlands

³Mechatronic Design

Delft University of Technology

Delft, the Netherlands

INTRODUCTION

One of the future challenges in the design of semi-conductor manufacturing equipment is the introduction of 450 mm wafers. In these machines the wafer is generally held and positioned on a wafer chuck. The position of the wafer chuck is controlled in an active position feedback loop, where it is assumed the wafer chuck behaves as a rigid body. As the wafer area increases so does the chuck's surface. To limit internal flexibility and keep internal eigen-frequencies at high levels, the chuck's thickness has to be increased as well. This leads to a significantly increased mass. In combination with the continuous demand for higher throughput this results in enormous acceleration forces. High acceleration forces risk perturbing accurate machine components, and lead to large power consumption and heat dissipation.

An alternative is to deal with the chuck's flexible modes in the position control loop directly, so called modal control [1]. In modal control the number of actuators (over-actuation) and sensors (over-sensing) is increased with respect to the rigid body Degrees of Freedom (Dof). Combined with high order controller the internal stiffness of the flexible chuck can be improved through control technical means. This method leads to additional control complexity and added component costs. In this paper a third alternative is presented where through mechanical topology optimization and careful actuator placement, an over-actuated flexible 450 mm wafer chuck is realized with high levels of performance, only 1 additional actuator and no additional control complexity.

TOPOLOGY OPTIMIZATION OF A 450 MM WAFER CHUCK

In mechatronics it is well known that actuator placement has a significant impact on the con-

trol performance of the positioning system. The effect of actuator placement on the excitation of mode shapes is illustrated in the first subsection. In the subsequent part a topology optimization for the 450 mm wafer chuck is presented exploiting the effects of actuator placement on modal actuation.

Influence of actuator placement on control

The dynamics of a free mass can be expressed in modal coordinates [1], where the overall displacement at the location of a sensor, $D(s)$, is formed by a combination of rigid body modes and flexible modes, Equation 1.

$$\frac{D(s)}{F(s)} = \frac{b_1 \cdot a_1}{m_1 s^2} + \sum_{n=2}^{\infty} \frac{b_n \cdot a_n}{m_n s^2 + c_n s + k_n} \quad (1)$$

m_n : Modal mass, 1 indicates rigid body mode

k_n : Modal stiffness

c_n : Modal damping

a_n : Modal actuator contribution

b_n : Modal sensor contribution

F : Force

The actuator contribution, a_n , is proportional to the mode shape at the actuator location [2]. When the actuator is positioned at the node of a flexible mode the actuator contribution for this specific mode will be equal to zero. When using multiple actuators with opposite modal contributions, the specific mode is not excited. Graphically this effect of actuator placement is illustrated in Figure 1.

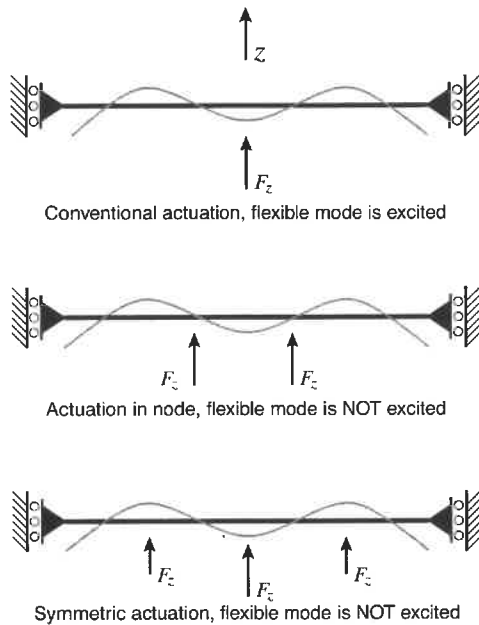


FIGURE 1. Effect of actuator location on excitation of a flexible mode.

By positioning the actuators at the right locations, excitation of specific flexible modes can be significantly reduced. This allows actuation of large flexible structures without exciting critical flexible modes. However, in mechanical structures there exist infinitely many flexible modes and node locations. To limit the amount of actuators, the mechanical structure can be optimized such that node locations of several modes coincide, limiting the number of required actuators. Such an optimization was performed for the 450 mm wafer chuck.

Topology optimization of a 450 mm chuck

Using a topology optimization, the mechanical structure of a wafer chuck was optimized such that several critical flexible modes are not excited while using only a limited number of additional actuators [3]. The optimization for the 450 mm wafer chuck was focused on the directions perpendicular to the plane of the wafer. These directions contain the modes with the lowest eigen-frequencies. In Figure 2 the general layout of the wafer chuck is shown.

A hole with varying diameter, d , parameter is cut out of the chuck and modifies the mode shapes. Simultaneously, the location of 4 actuators is varied along the diagonals of the chuck to search for the optimal actuator location for each hole diameter. The optimization process is illustrated using

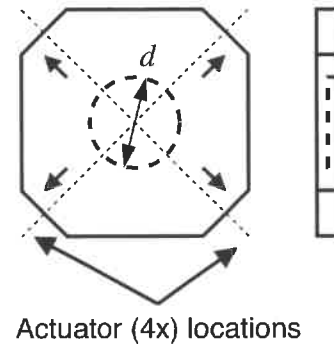


FIGURE 2. Wafer chuck of a 450 mm wafer stage. The diameter of the hole, d , is varied to modify the mode shapes. The actuators are displaced along the diagonals until the optimal actuator location is found.

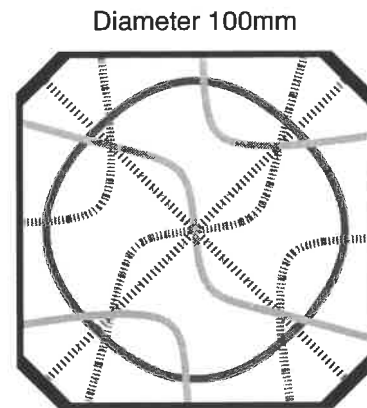


FIGURE 3. Resulting nodal lines of 4 flexible modes with a hole diameter of 100 mm. The nodal lines coincide on different locations.

the nodal lines of three flexible modes in Figure 3 and 4. At the optimum diameter of 446 mm the nodal lines of the 4 flexible modes coincide on a single point which also is the optimal actuator location. Figure 5 shows the mode shapes taken into account during the optimization process.

MECHANICAL DESIGN

A demonstrator for studying over-actuation on a 450 mm wafer chuck was realized at MI-Partners. The mechanical layout of the over-actuated 450 mm chuck is shown in Figure 6. The chuck weighs 7 kg and has a thickness of 20 mm. The design allows both conventional actuation on the 3 Bessel's points and over-actuation with 4 actuators. This allows performance of both methods to be compared.

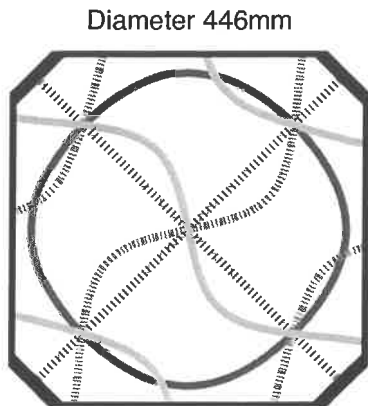


FIGURE 4. Resulting nodal lines of 4 flexible modes with a hole diameter of 446 mm. The nodal lines coincide at one point, indicating the optimal actuator location.

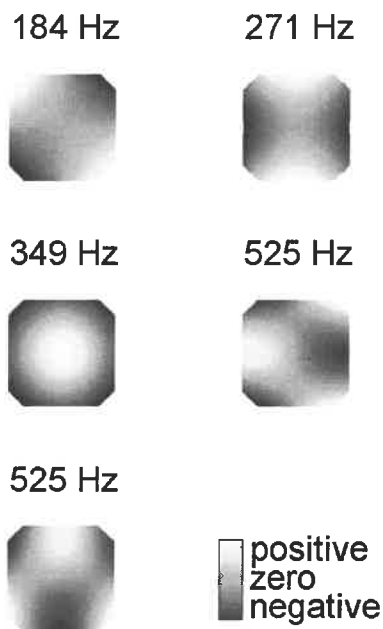


FIGURE 5. First 5 flexible mode shapes of the 450 mm wafer chuck. These mode shapes were taken into account during the optimization process.

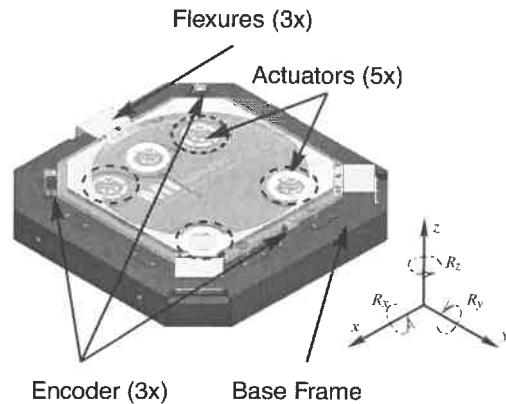


FIGURE 6. Mechanical layout of over-actuated chuck. The chuck can be actuated using the 4 actuators on the optimal actuator locations (encircled with black dashed line); or the chuck can be actuated by conventional actuation in the 3 Bessel's points (encircled with white line).

The optimization was focused on the critical modes in the plane perpendicular to the wafer (z , R_x , R_y). The other DoF are constrained using mechanical flexures. In an actual system the flexures can easily be replaced by an active control loop since in these directions the chuck behaves as a rigid body. The position of the chuck is measured using 3 linear encoders. The system uses "moving magnet actuators" developed by Magnetic Innovations. To ensure the flexible modes of the base frame are not excited by the actuators, the coils are attached to low frequency suspended masses within the frame.

CONTROL STRATEGY

Figure 7 shows the conventional control strategy and the over-actuated control strategy applied to the chuck. The chuck uses 3 position sensors (equal to the DoF) to reconstruct the global center of gravity coordinates, namely z , R_x and R_y . Upon these coordinates conventional PID controllers with notches are designed. The output of the controller are the desired center of gravity forces, F_z , M_x and M_y . A static actuation matrix distributes the forces over the different actuators. The difference between the conventional strategy and over-actuated strategy is a difference in static actuation matrix.

PERFORMANCE

In Figure 8 the difference between measured open-loop frequency response of the conventional actuated chuck and the over-actuated

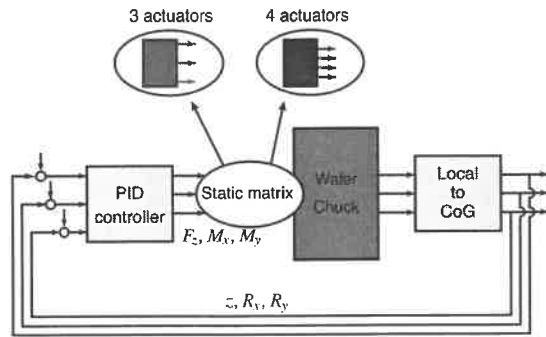


FIGURE 7. Control strategy of the wafer chuck. The controlled variables are the center of gravity coordinates of the chuck. The difference in control strategy between over-actuation and conventional actuation lies in the different static actuator matrices.

chuck is shown. The conventionally actuated chuck has a -2 -4 characteristic caused by the flexible mode at 180 Hz. In the over-actuated situation the chuck shows a -2 slope up to 800 Hz, allowing a significant improvement in control performance.

In Table 1 the attained bandwidths and servo-errors of the wafer chuck under conventional actuation and over-actuation are shown. In both cases the controller has been optimized to reach high bandwidth and positioning performance using standard PID controllers with 2 to 3 notches. The main disturbance source acting on the chuck originated in floor vibrations. By using over-actuation an improvement of a factor 2 in bandwidth was attained and a factor 5 in disturbance rejection.

TABLE 1. Servo-error of Conventional and Over-actuated chuck

	Conventional	Over-actuated
Bandwidth		
z	50 Hz	100 Hz
R_x	60 Hz	100 Hz
R_y	50 Hz	100 Hz
Servo-error		
z	30 nm	7 nm
R_x	150 nrad	35 nrad
R_y	135 nrad	30 nrad

CONCLUSION

The paper has demonstrated that by using over-actuation on a large flexible structure, a significant step in controlled performance can be made.

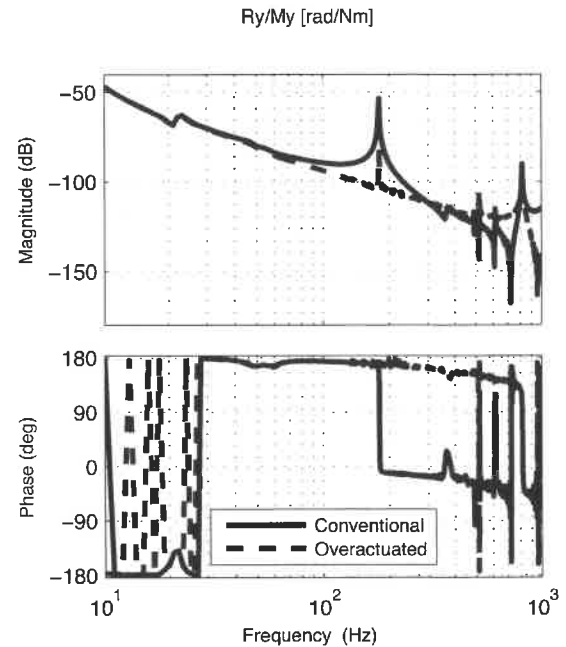


FIGURE 8. The measured frequency response in R_y direction, showing that the over-actuated system outperforms the conventionally actuated system. Note that the conventional system has a -180° phase shift.

This was achieved with only 1 additional actuator and limited additional control complexity. The result was achieved by topology optimization of the wafer chuck's mechanics and careful actuator placement.

ACKNOWLEDGMENTS

The authors thank Magnetic Innovations and Delft University of Technology for their cooperation. SenterNovem and SRE have financially supported the research.

REFERENCES

- [1] Gawronski W. Dynamics and control of structures: a modal approach. New York: Springer-Verlag; 1998.
- [2] Gawronski W. Modal actuators and sensors. Journal of Sound and Vibration. 2000;229(4):1013–1022.
- [3] Langelaar M. Exploratory Investigation into Topology Optimization of Overactuated Systems. Delft University of Technology; 2009.

A DESKTOP ELECTROHYDRODYNAMIC JET PRINTING SYSTEM WITH INTEGRATED HIGH-RESOLUTION SENSING AND CONTROL

Sandipan Mishra, Kira Barton, and Andrew Alleyne
Mechanical Science and Engineering Department,
University of Illinois, Urbana, IL 61801 USA

Abstract

This paper discusses the design and integration of a desktop system with sensing and control capability for Electrohydrodynamic jet (E-jet) printing. E-jet printing is a micro/nano-manufacturing process that uses an electric field to induce fluid jet printing through micro/nano-scale nozzles. This enables better control and resolution than traditional jet-printing processes. The printing process is predominantly controlled by changing the voltage potential between the nozzle and the substrate. The push to drive E-jet printing towards a viable micro/nano-manufacturing process has led to the design of a compact, cost effective, and user friendly desktop E-jet printing system. The hardware and software components of the desktop system are described in the paper. A current detection system is designed for monitoring the printing. We also propose a two DOF feedback-feedforward control law for control of the printing process. Experimental results are presented to validate the performance of the desktop system.

INTRODUCTION

As the demand for micro- and nano-scale devices in electronics, biotechnology and micro-electromechanical systems has increased, efforts have been made to adapt current graphic art printing techniques to address this need. Conventional methods for graphic art printing such as inkjet printing include applying heat to induce a vapor bubble to form and eject a droplet of ink through a nozzle, and piezoelectric printers which squeeze a glass tube to eject ink [1]. The minimum printing resolution that can be created reliably for these methods ranges from 20-30 μ m. This coarse resolution is due to a combination of nozzle sizes and droplet placement. Smaller nozzle sizes may become clogged due to the ink viscosity, while the vibrations caused by the piezoelectric actuators often lead to variations in the droplet placement [2]. These traditional graphic art approaches cannot be used for high-resolution manufacturing due to size and accuracy limitations.

Electrohydrodynamic jet (E-jet) printing is a technique that uses electric fields to create fluid flow necessary to deliver ink to a substrate for high-resolution (< 10 μ m) patterning applications [3]. E-jet has been gaining momentum in the past few years as a viable printing technique, especially in the micro- and nano-scale range [4, 5, 6]. As the advantages of E-jet printing become more apparent (e.g. the potential for purely additive operations, the ability to directly pattern biological materials for biosensors, drop-on-demand functionality for chemical mixing and sensor fabrication, and high-resolution printing for printed electronics), the necessity for compact, affordable, and user friendly E-jet printing systems increases.

Our group has previously demonstrated high resolution E-jet printing [3] using expensive custom-built equipment. In this paper, we present a desktop system for E-jet printing, designed from commercial off-the-shelf technology (COTS) components, competitive in terms of cost with many of the commercially-available printers but capable of much higher resolutions. The system consists of the necessary hardware and software for standard E-jet printing as well as sensing and control of the printing process. More specifically, this paper will focus on 1) the design and fabrication of a micro/nano-manufacturing testbed for E-jet printing, and 2) the development of a sensing and control system for E-jet printing.

ELECTROHYDRODYNAMIC JET PRINTING

E-jet printing uses electric field induced fluid flows through microcapillary nozzles to create devices in the micro/nano-scale range. The patterning of wide ranging classes of inks in diverse geometries, as well as printed examples of functional circuits and sensors demonstrating the diverse applications of E-jet printing are provided in [3].

Figure 1 presents a schematic of the E-jet printing process. The main elements for E-jet printing include an ink chamber, controlled pressure supply, glass nozzle tip, substrate, and positioning system. The printing conditions are controlled

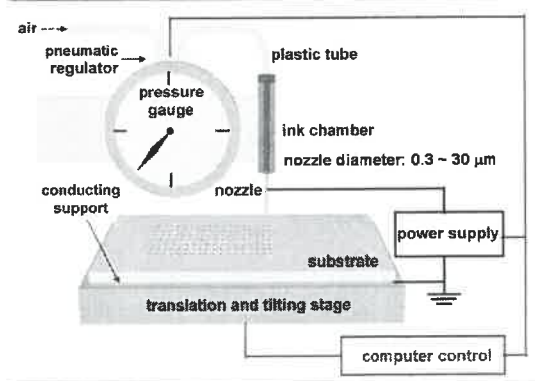


FIGURE 1. Schematic of the E-jet printing set-up. Taken from [3].

through the back pressure (air applied to the nozzle), the offset height, and the applied voltage potential between a conducting nozzle tip and substrate. Changes in back pressure, stand-off height, and applied voltage affect the size and frequency of the droplets. These changes result in different jetting modes (e.g. pulsating, stable jet, e-spray) which can be used to achieve various printing requirements. Choi *et al* [7] proposed the following relationship for frequency of jetting f with the voltage potential V and standoff height h : $f = K \left(\frac{V}{h} \right)^{3/2}$, where K is a scaling constant dependent on the viscosity of the ink, the nozzle diameter, applied back pressure, and permittivity of free space. For a detailed derivation of this, the interested reader is referred to [7].

For E-jet printing, an applied voltage potential is generated between a conducting nozzle and substrate. Note that the nozzle tip and substrate are generally coated with metal to ensure conductivity. Additionally, if the surface of the desired substrate is nonconductive, one can use a conductive layer under a nonconductive substrate provided that the thickness of the nonconductive substrate is within a certain range. A voltage applied to the nozzle tip causes mobile ions in the ink to accumulate near the surface at the tip of the nozzle. The mutual Coulombic repulsion between the ions introduces a tangential stress on the liquid surface that, along with the electrostatic attraction to the substrate, deforms the meniscus into a conical shape (called the Taylor cone after Sir Geoffrey Ingram Taylor who first reported it in 1964) as described in [3]. At some point, the electrostatic stress overpowers the surface tension and droplets eject from the cone. Figure 2 illustrates the change in the apex of the ink meniscus due to

an increase in voltage.



FIGURE 2. Illustration of the change in the meniscus of the fluid due to an increase in voltage potential between the nozzle tip and the substrate. Taken from [3].

The pinching off of the fluid from the apex of the cone results in droplets that are typically smaller than the nozzle (micro-pipette) diameter. In an effort to package and simplify the process and make E-jet printing more accessible to researchers working on potential printing applications in micro/nano-manufacturing, a desktop printing system has been developed. Details describing the necessary hardware for this system are provided in the following section.

HARDWARE AND SOFTWARE INTERFACING

From the previous section, the hardware requirements for the desktop E-jet system consist of: the positioning elements, the pressure and vacuum pumps, the visualization system, the toolbit and substrate mounts, the electrical connections for generating the required voltage potential, and the housing elements.

As can be seen from Figure 3, the positioning system consists of x- and y-axis electronic positioning stages, a manual z-axis, and a manual rotary axis. Manual z and rotary axes were used to minimize costs. Back pressure and voltage potential compensate for any height irregularities using the relationship provided in the earlier section. The

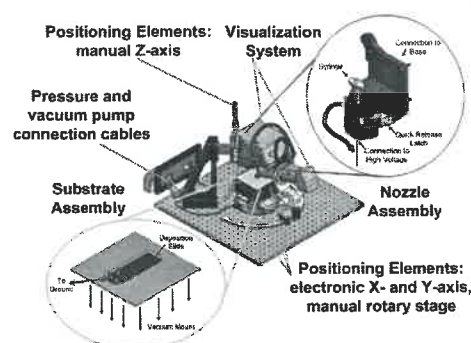


FIGURE 3. Desktop E-jet system with specific hardware requirements identified. Note that the major positioning and jetting components for the desktop E-jet system are sized to fit a typical lab desktop.

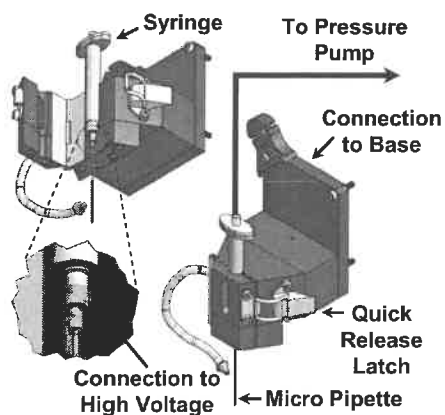


FIGURE 4. Nozzle mount for the E-jet process. Note the electrical connection used to apply a high voltage signal to the treated micropipette.

pressure pump applies back pressure to the syringe, while the vacuum pump is used to attach the substrate to the substrate mount. The visualization system includes a high-resolution camera and magnification lens mounted to a 180 degree rotary track, as well as a fiber optic light with adjustable arms. The housing is made up of a breadboard and glass enclosure (shown in Fig. 11). All of the items described thus far have been purchased as off-the-shelf components from vendors.

The remaining hardware consists of components that are specific to the E-jet printing process. The toolbit and substrate mounts and the electrical connections residing within these components are critical to the E-jet process and require custom designs. Figure 4 illustrates one of the toolbit mounts. This mount is designed for single nozzle deposition. An off-the-shelf syringe containing the deposition ink is connected to the pressure pump and a Luer lock micro-pipette ranging in tip size from 300 nm to 10 microns. The micro-pipette (nozzle) is sputter coated with metal prior to assembly to ensure an electrical connection along the length of the nozzle. Additionally, the pipette tip is treated with a hydrophobic coating to minimize wicking of the ink along the nozzle. The conductive base of the pipette makes an electrical connection with the mount using contact pins.

The interfacing of the desktop system through LabVIEW was designed to integrate the two major subsystems: (a) the positioning system (linear motors and the motor drivers) and (b) the electri-

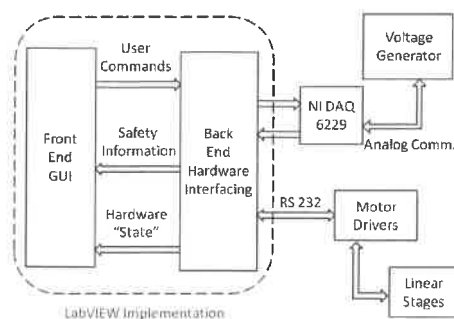


FIGURE 5. Desktop E-jet System Software-Hardware Interface

cal system (high voltage amplifier). LabVIEW was chosen for software interfacing due to its easy to use front end graphical interface and the accessibility and modular capabilities of its back end platform. There are two modes of operation for the software. In manual mode, the user has control over position and voltage signals. This mode is used to test the E-jet process for determining suitable voltages for consistent jetting conditions. In the automated operation mode, a set of pre-programmed commands can be loaded and executed sequentially to generate a specific pattern on the substrate through coordination of the voltage and position commands. The voltage commands, however, can be over-written by the user while in the automated operation mode.

Figure 5 shows a schematic of the software-hardware interfacing. The voltage amplifier is controlled and monitored through analog communication via an NI-6229 DAQ board. On the other hand, the motor drivers are controlled over a serial port (RS 232) communication link. The front end GUI enables the user to monitor safety signals and send control signals for operation over these communication links.

CURRENT DETECTION

Traditionally, the E-jet process has been monitored primarily through imaging, both online and offline. A camera is used to view the emission of the droplet from the nozzle onto the substrate. However, there are some significant disadvantages to this monitoring method. Firstly, image processing is time consuming and is unsuitable for feedback control of the process with low computation power. Further, without advanced image processing algorithms, this monitoring method necessitates the presence of a human operator for supervision. In order to address both these issues, this paper proposes the use of a current

detection system for sensing process operating conditions for E-jet printing. This current detection system is better suited for online monitoring and automated control of the E-jet process since the measurement and data analysis are simple and can be done at the same time-scale as the process (up to 1 kHz).

Current-detection based process characterization of the process is based on the following fundamental physical phenomenon during E-jet. When a charged droplet is released from the nozzle, the voltage source generates a small current to neutralize the imbalance in charge in the fluid inside the nozzle. By detecting this current, the time of droplet release can be determined. This measurement scheme is termed *Nozzle-side* measurement. An alternate scheme measures the current discharged through the substrate. When a charged droplet from the nozzle hits the conductive substrate, the charge is dissipated through to the ground. This current can be measured by connecting a current sensor to the substrate-ground connection. This measurement scheme is termed *Substrate-side* measurement. Figure 6 shows a schematic of the nozzle-side current measurement setup. The high voltage source is connected to the substrate side, while a current sensor is connected to the nozzle side. The free end of the current sensor drains to ground.

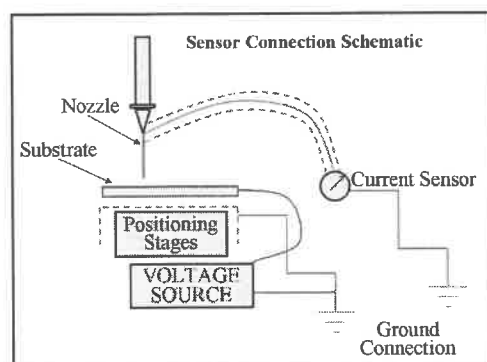


FIGURE 6. Schematic of the nozzle-side current measurement setup for the E-jet process.

The frequency of jetting can be determined by measuring the time elapsed between two successive jets. Each peak in the current signal corresponds to a single jet. This is illustrated below in Figure 7. This signal can then be used in the control algorithm for regulation of frequency about a set point.

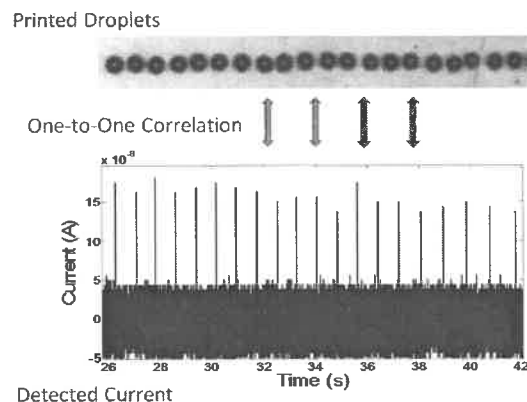


FIGURE 7. Illustration of the one-to-one correlation between the printed droplets and the measured current peaks.

The detailed plot of the current peak when a single droplet is released is shown in Figure 8. The peak current is proportional to the size of the droplet (dependent on the applied voltage and back pressure). This makes intuitive sense since a larger droplet carries more charge. The duration of the jet is also directly proportional to the size of the droplet. The peak current is typically of the order of 10-100s of nanoamperes (in this case: 520nA). These small currents necessitate very high quality shielding and noise suppression. The signal to noise ratios are typically of the order of 5 – 10. Further, the duration of the jet is generally less than 50 μ s (in this case: 30 μ s).

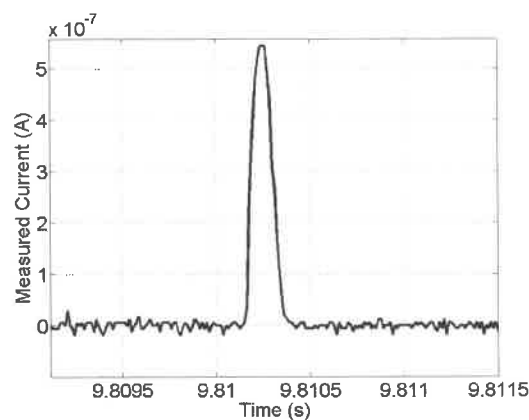


FIGURE 8. Detailed image of a current peak corresponding to a single released droplet. The current peak has an amplitude of 520nA and a duration of 30 μ s

CONTROL SYSTEM DESIGN

The consistency of droplet deposition, i.e. the jetting frequency, is a key metric for evaluation of the

E-jet printing process. The controllable input signal is the applied voltage difference between the nozzle and the substrate. In open loop operation of the process, a fixed voltage difference is applied to the nozzle and the substrate. However, this strategy results in substantial variation of jetting frequency because process parameters such as standoff height and wetting properties of the nozzle are subject to variation during the course of the printing process. In order to overcome this, we propose the use of a two DOF feedback and feedforward control algorithm to regulate the jetting frequency.

The *feedback* controller is an integral control law of the form

$$V_{fb}(k+1) = V_{fb}(k) + K_i (f_{des} - f(k)) \quad (1)$$

where K_i is the integral control gain, and $f(k)$ is the measured jetting frequency. Notice that the index k refers to the sample instant; however, $f(k)$ is *not* updated at every sample instant. $f(k)$ is updated only when a jet is detected.

In order to find the ideal feedforward voltage profile to pre-compensate for change in substrate standoff height, we implement an ILC algorithm for updating the *feedforward* voltage signal based on jetting frequency estimates from the previous runs of the process. The optimized feedforward control signal profile V_{ff} shown in 9 is obtained by running the learning algorithm to convergence within a bound.

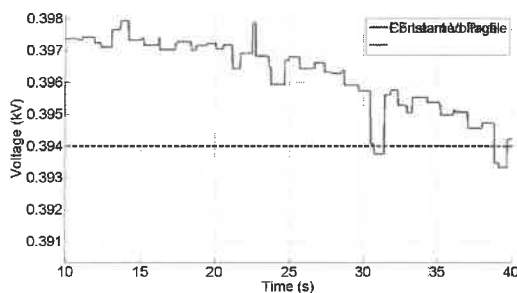


FIGURE 9. Learned Feedforward Voltage Profile V_{ff} .

Once the optimized feedforward signal has been identified, it can be included in the total voltage input signal along with feedback control. This 2-DOF control law is given by

$$V_{tot}(k) = V_{ff}(k) + V_{fb}(k). \quad (2)$$

The feedforward signal acts as the baseline voltage profile, while the feedback signal acts as supplemental control to minimize short-term stochastic disturbances. The addition of the feedforward signal decreases the feedback gain required to optimize the jetting frequency since the large voltage increases due to the offset height are taken care of by the feedforward signal. Figure 10 shows the schematic of the plant and 2-DOF control system.

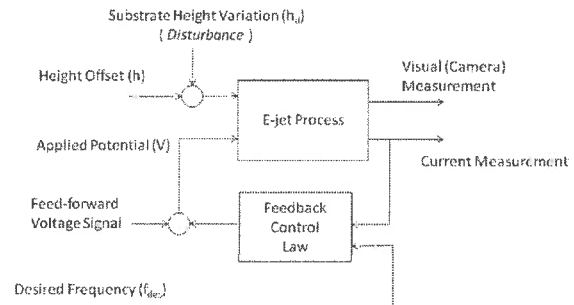


FIGURE 10. Block diagram of the E-jet printing process with feedback and feedforward control.

EXPERIMENTAL RESULTS

In order to validate the performance capabilities of the desktop E-jet printing system (Fig. 11), a sample image was drawn using the process diagrammed in Fig. 12. Operating from the manual mode on the GUI, an initial calibration was performed to determine suitable XY position, z-axis offset height, back pressure and voltage input for a desired jetting frequency. Switching over to the automated mode, a series of position and voltage commands were uploaded into the GUI. Sequential implementation of the uploaded commands resulted in a block 'I' image shown in Fig. 13.



FIGURE 11. Electrohydrodynamic jet printing system designed to minimize cost (< \$50,000), system dimensions, and set-up time, while maintaining high resolution E-jet printing capabilities.

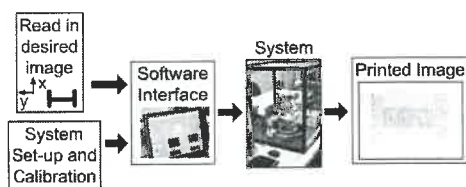


FIGURE 12. Process diagram of the E-jet printing system.

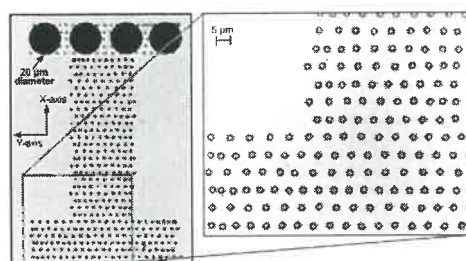


FIGURE 13. Block "I" printed using the desktop E-jet system. Image was printed from a nozzle diameter of $5\text{ }\mu\text{m}$ resulting in printed droplets with an average measured diameter of $2.8\text{ }\mu\text{m}$. Typical ink jet droplets with a $20\text{ }\mu\text{m}$ diameter have been superimposed on the printed image for comparison purposes.

Using a $5\text{ }\mu\text{m}$ nozzle tip (micro-pipette), the desktop system printed droplets with an average measured diameter of $2.8\text{ }\mu\text{m}$. The droplet size is correlated to several process variables including: nozzle tip, ink viscosity, offset height, back pressure, and applied voltage potential between the conducting nozzle tip and substrate. Changes in these conditions will result in variations in the droplet diameter and jetting frequency. For the system in Fig. 11, the process variables were shown to be consistent over a printing area of $5\text{ mm} \times 5\text{ mm}$, thereby indicating minimal built-in tilt offset with the printer. The block 'I' was printed by rastering back and forth along the y-axis with a fixed jetting voltage determined during the initial calibration. The natural pulsating jet mode of the meniscus resulted in slight discrepancies in droplet placement. For droplet size comparison, droplets representing a typical ink jet printing resolution of approximately $20\text{ }\mu\text{m}$ have been superimposed on the E-jet printed image in Fig. 13. These results clearly indicate the ability of E-jet printing to surpass the printing resolution of typical ink jet printers.

CONCLUSION

The availability of compact, affordable, and user friendly test platforms for mi-

cro/nanomanufacturing processes is a critical part of enabling the transition of these processes into mainstream manufacturing systems. The major challenge is providing affordable test platforms for researchers to further develop the process and associated applications. E-jet printing is an emerging manufacturing technology that has potential in widespread applications. This paper presented a small and affordable desktop system for E-jet printing with integrated sensing and control capability.

REFERENCES

- [1] Calvert P. Inkjet Printing for Materials and Devices. *Chem Mater.* 2001;13(10):3299–3305.
- [2] Szczech J, Megaridis C, Gamota D, Zhang J. Fine-Line Conductor Manufacturing Using Drop-on-Demand PZT Printing Technology. *IEEE Transactions on Electronics Packaging Manufacturing.* 2002;25(1):26–33.
- [3] Park JU, Hardy M, Kang SJ, Barton K, Adair K, Mukhopadhyay D, et al. High-resolution Electrohydrodynamic jet printing. *Nature Materials.* 2007 August;6:782 – 789.
- [4] Jayasinghe S, Qureshi Q, Eagles P. Electrohydrodynamic Jet Processing: An Advanced Electric-Field-Driven Jetting Phenomenon for Processing Living Cells. *Small.* 2006;2:216–219.
- [5] Youn DH, Kim SH, Yang YS, Lim SC, Kim SJ, Ahn SH, et al. Electrohydrodynamic micropatterning of silver ink using near-field electrohydrodynamic jet printing with tilted-outlet nozzle. *Applied Physics A.* 2009;96:933–938.
- [6] Wang K, Paine M, Stark J. Fully voltage-controlled electrohydrodynamic jet printing of conductive silver tracks with a sub $100\text{ }\mu\text{m}$ linewidth. *Journal of Applied Physics.* 2009;106:0249071 – 0249074.
- [7] Choi HK, Park JU, Park OO, Ferreira PM, Georgiadis JG, Rogers JA. Scaling laws for jet pulsations associated with high-resolution electrohydrodynamic printing. *Applied Physics Letters.* 2008;92(12):123109.

Author Index

Abbott, Ben	75	Madariaga, Jon	59
Aarden, W. W. J.	67	Manske, Eberhard	40
Al-Bender, Farid	117	Mason, Ken	75
Alleyne, Andrew G.	145	Matichard, Fabrice	75
Amin-Shahidi, Darya	51	McCarthy, Kevin J.	135
Azpeitia, José Luis	59	Merz, Michael	17
Baggen, Mark*	66	Mishra, Sandipan	145
Barnum, Sam	75	Mittleman, Richard	75
Barton, Kira L.	145	Mock, Christopher D.	87
Boshuisen, Rein	141	Norg, Meindert L.	29
Brecher, Christian	17	Nuij, Pieter W. J. M.	129
Bringmann, Bernhard	45	O'Reilly, Brian	75
Brinken, Frank	45	Overcash, Jerald L.	51
Bristow, Douglas*	65	Panas, Robert M.	105
Chong, Shin-Horng	123	Percle, Brandon D.	40
Coyne, Dennis	75	Profeta, Joseph A.	55
Cullinan, Michael A.	105	Radkins, Hugh	75
Culpepper, Martin L.	105	Ramet, Céline	75
De Bra, Daniel B.	75	Reck, Matthew L.	35
Dorozhovets, Nataliya	40	Rijlaarsdam, David J.	129
Dow, Thomas A.	96	Rodríguez de Yurre, Juan C.	59
Eijsinger, Jamie	67	Ruben, Shalom D.	111
Erkorkmaz, Kaan	87	Ruiz de Argandoña, Ismael	59
Foley, Stephany	75	Sato, Kaiji	123
Garrard, Kenneth P.	96	Schoukens, Johan	129
Giaime, Joseph A.	75	Shapiro, Brett N.	69
Gordon, Daniel J.	87	Shoemaker, David	75
Gray, Cory	75	Smeds, Kristofer	23
Hanson, Joe	75	Sohn, Alex	96
Hausotte, Tino	40	Stein, Andy	75
Heertjes, Marcel F.	81	Steinbuch, Maarten	81, 129
Hocken, Robert J.	51	Trumper, David L.	51
Hunter, Bradley	135	Tsao, Tsu-Chin	111
Jäger, Gerd	40	Uriarte, Luis G.	59
Kissel, Jeff	75	Usman, Irfan-ur-rab	23
Lantz, Brian T.	75	van den Berg, C. J. M.	67
Laro, Dick A. H.	67, 141	van Eijk, Jan	67, 102, 141
Lindemann, Dominik M.	17	van Engelen, Arjan J. P.	81
Ljubicic, Dean M.	51	van Geffen, Vincent	129
Lu, Xiaodong	23	Vervoordeldonk, Michiel J*	66
Ludwick, Stephen J.	55	Wenzel, Christian	17
Machleidt, Torsten	40	Wilson, Curtis S.	11
MacInnis, Myron	75		

* Extended abstract unavailable at press time

ASPE 2010 Sustaining Corporate Sponsors



3M Company
St. Paul, MN 55144
Phone: (651) 737-3790
www.mmm.com



ADCOLE Corporation
669 Forest Street
Marlborough, MA 01752
Phone: (508) 485-9100
Fax: (508) 481-6142
www.adcole.com



Aerotech, Inc.
101 Zeta Drive
Pittsburgh, PA 15238-2897
Phone: (412) 963-7470
Fax: (412) 963-7459
www.aerotech.com



Ametek Precitech, Inc.
44 Blackbrook Road
Keene, NH 03431
Phone: (603) 357-2511
Fax: (603) 358-6174
www.precitech.com



Chardon Tool
115 Parker Court
Chardon, OH 44024
Phone: (440) 286-6440
Fax: (440) 286-7165
www.ChardonTool.com



Equilibria
472 Amherst Street
Unit 25
Nashua, NH 03063
Phone: (603) 880-7722
www.eqlib.com



Fabreeka International, Inc.
P.O. Box 210
1023 Turnpike Street
Stoughton, MA 02072
Phone: (781) 341-3655
Toll-free: (800)-322-7352
Fax: (781) 341-3983
www.fabreeka.com



K & Y Diamond
2645 Diab
St. Laurent Quebec H4S 1E7
CANADA
Phone: (514) 333-5606
Fax: (514) 339-5493
www.kydiamond.ca



Lion Precision
563 Shoreview Park Road
St. Paul, MN 55126-7014
Phone: (651) 484-6544
Fax: (651) 484-6824
www.lionprecision.com

ASPE 2010 Sustaining Corporate Sponsors



Makino
7680 Innovation Way
Mason, OH 45040
Phone: (513) 573-7200
Fax: (513) 573-7360
www.makino.com



Mitutoyo America Corporation
945 Corporate Boulevard
Aurora, IL 60504-9102
Phone: (630) 723-3619
Fax: (630) 978-6471
www.mitutoyo.com



Moore Nanotechnology Systems LLC
426A Winchester Street
Keene, NH 03431-0605
Phone: (603) 352-3030
www.nanotechsys.com



Piezo Nano Positioning
PI (Physik Instrumente) L.P.
16 Albert Street
Auburn, MA 01501
Phone: (508) 832-3456
www.physikinstrumente.com



Professional Instruments Company
7800 Powell Road
Hopkins, MN 55343
Phone: (952) 933-1222
Fax: (952) 933-3315
www.airbearings.com



Specialty Components, Inc.
14 Village Lane
Wallingford, CT 06492
Phone: (203) 284-9112
Fax: (203) 672-0371
www.specialtycomponents.com



Trust Automation, Inc.
143 Suburban Road
San Luis Obispo, CA 93401
Phone: (805) 544-0761
Fax: (805) 544-4621
www.trustautomation.com



Veeco
2650 East Elvira Road
Tucson, AZ 85706-7123
Phone: (520) 741-1044
www.veeco.com



Zygo Corporation
Laurel Brook Road
Middlefield, CT 06455-0448
Phone: (860) 347-8506
Fax: (860) 347-8372
Toll-free: (800) ZYGO-NOW
www.zygo.com

ASPE 2010 Corporate Sponsors

Contour Fine Tooling, Inc.

143 Jaffrey Road
Marlborough, NH 03455
Phone: (603) 876-4908
www.contour-diamonds.com

Corning Inc.

Science & Technology Division
HP-ME-03-055
Corning, NY 14831
Phone: (607) 974-3842
www.corning.com

Cranfield Precision

Division of Cinetic Landis Lund, Ltd.
Woburn House, 3 Adams Close
Kempston, Bedford MK42 7JE
United Kingdom
Phone: +44-1234-312820
www.cranfieldprecision.com

euspen Ltd.

Building 70
Cranfield University
Bedford, MK43 0AL, United Kingdom
Phone: +44-1234-754024
www.euspen.org

Moog, Inc.

30 Jamison Road
East Aurora, NY 14052-0018
Phone: (716) 687-7764
www.moog.com/industrial

Nano Corporation

5-12-5 Kugahara, Ota-ku
Tokyo 144-0042, Japan
Phone: +81-3-5735-6631
www.nanowave.com

nanoPrecision Products, Inc.

411-B Coral Circle
El Segundo, CA 90245
Phone: (310) 773-3550
www.nanoprecisionproducts.com